

LONDON
SCHOOL of
HYGIENE
& TROPICAL
MEDICINE



LSHTM Research Online

McElvenny, DM; (2017) Meta-analysis of Rare Diseases in Occupational Epidemiology. PhD thesis, London School of Hygiene & Tropical Medicine. DOI: <https://doi.org/10.17037/PUBS.03894558>

Downloaded from: <https://researchonline.lshtm.ac.uk/id/eprint/3894558/>

DOI: <https://doi.org/10.17037/PUBS.03894558>

Usage Guidelines:

Please refer to usage guidelines at <https://researchonline.lshtm.ac.uk/policies.html> or alternatively contact researchonline@lshtm.ac.uk.

Copyright \copyright\ and Moral Rights for the papers on this site are retained by the individual authors and/or other copyright owners

<https://researchonline.lshtm.ac.uk>

LONDON
SCHOOL *of*
HYGIENE
& TROPICAL
MEDICINE



Meta-analysis of Rare Diseases in Occupational Epidemiology

DAMIEN MARTIN MCELVENNY

Thesis submitted in accordance with the requirements for the degree of

Doctor of Philosophy of the University of London
JANUARY 2017

Department of Social and Environmental Health Research

Faculty of Public Health and Policy

LONDON SCHOOL OF HYGIENE & TROPICAL MEDICINE

Funded by the Health & Safety Executive

ACKNOWLEDGEMENTS

I would like to thank my wife Jackie my son Daniel and daughter Kate for their continued support and forbearance during the completion of this work.

I would also like to thank my late parents Francis and Margaret McElvenny for their encouragement to pursue a university education.

I would like to acknowledge the funding of my former employer The Health and Safety Executive who paid my tuition fees and in particular Dr. John Osman for allowing me and encouraging me to undertake this work.

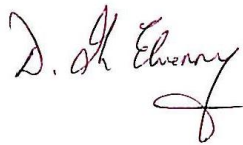
Finally, I would like to thank my supervisor for his expert guidance and his endless patience.

DECLARATION

I, Damien Martin McElvenny, confirm that the work presented in this thesis is my own.

Where information has been derived from other sources, I confirm that this has been indicated in the thesis.

Signed

A handwritten signature in black ink, reading "D. M. McElvenny". The signature is written in a cursive style with a large, stylized 'M' and a long, sweeping underline.

January 2017

ABSTRACT

At the outset of this research in the early 2000s, the application of meta-analysis in observational epidemiology, including occupational epidemiology was regarded as controversial because of the greater potential for bias in such studies compared with randomized controlled clinical trials. This remains true even in 2017.

The overall aim for this research is to identify the best approaches to the use of meta-analysis in the occupational health setting, and to summarise through meta-analysis the evidence as to whether or not occupational exposure to formaldehyde is an occupational carcinogen.

Chapter 2 of this thesis concluded that meta-analysis in occupational epidemiology was becoming increasingly popular, but that its limitations appear not to have been heeded in practice. Two principal issues were identified: the heterogeneity of exposures estimates and the pooling of standardized mortality ratios from different study populations with different characteristics including length of follow-up. By 2016, neither of these issues had been addressed in the literature.

Chapter 3 contains a description of the published statistical methods available for meta-analysis up to 2001. Methodological developments continue within the discipline of randomized controlled clinical trials and recently include the advent of multivariate methods and network analysis. It has been recommended that any new methods should be backed up by simulations.

Chapter 4 concludes that the default approach adopted by most statistical packages could not deal with such studies and excluded them from any calculations. In meta-analyses of rare diseases, this biases the meta-relative risks upwards. Approaches that

avoided exclusion of such studies are considered, in particular analyses on the original untransformed scale rather than the log scale.

Chapter 5 concludes that there remains insufficient evidence for an association. There was significant heterogeneity in the lung cancer results and so this is a random effects analysis; the analyses for nasopharyngeal and sinonasal cancer contains no such heterogeneity and so are fixed effect analyses.

The final chapter concludes that, when studies with zero cases are not excluded, there is insufficient evidence of an increased risk of lung cancer, nasopharyngeal cancer or sinonasal cancer, and that further methodological developments are still required to deal with the pooling of occupational epidemiological studies in relation to study characteristics, exposure assessments and standardised mortality ratios.

CONTENTS

1	INTRODUCTION	10
1.1	THE META-ANALYSIS CONTROVERSY	10
1.2	THE ENVIRONMENT OF OCCUPATIONAL REGULATORY RISK ASSESSMENT AND LIMIT SETTING.....	10
1.3	THE AIMS OF THIS THESIS	11
1.4	THE OUTLINE OF THIS THESIS	11
2	LITERATURE REVIEW 1 – META-ANALYSIS: PAST PRACTICE IN OCCUPATIONAL EPIDEMIOLOGY	12
2.1	PUBLISHED REVIEW OF PAST PRACTICE IN OCCUPATIONAL EPIDEMIOLOGY.....	12
2.2	PAST PRACTICE IN OCCUPATIONAL EPIDEMIOLOGY: A BRIEF UPDATE	23
3	LITERATURE REVIEW 2 - META-ANALYSIS METHODS AVAILABLE FOR THE OCCUPATIONAL HEALTH SETTING.....	25
3.1	INTRODUCTION: SCOPE AND STRUCTURE OF THE REVIEW	25
3.2	SEARCH METHODOLOGY	25
3.3	METHODS FOR ESTIMATING MEAN MEASURE OF EFFECT	26
3.3.1	<i>Fixed effect methods.....</i>	26
3.3.2	<i>Random effects methods.....</i>	35
3.4	METHODS FOR ASSESSING THE PRESENCE OF BETWEEN-STUDY HETEROGENEITY ...	41
3.4.1	<i>Generic Cochran chi-squared test</i>	41
3.4.2	<i>SMRs</i>	42
3.4.3	<i>Peto method.....</i>	42
3.4.4	<i>Correlation coefficients.....</i>	42
3.5	METHODS FOR EXPLORING BETWEEN-STUDY HETEROGENEITY	43
3.5.1	<i>Graphical methods</i>	43
3.5.2	<i>Subgroup analysis – stratification by study characteristics</i>	44
3.5.3	<i>Fixed effect and random effects regression models</i>	45
3.6	PUBLICATION BIAS.....	46
3.6.1	<i>Funnel plots (identification of publication bias).....</i>	47
3.6.2	<i>The file-drawer method (measurement of extent of publication bias)</i>	47
3.6.3	<i>Brief overview of other methods for dealing with publication bias ...</i>	48
3.7	OTHER META-ANALYTICAL METHODS.....	49
3.7.1	<i>Methods for combining dose-response data</i>	49
3.7.2	<i>Methods for assessing and adjusting for study quality</i>	49
3.7.3	<i>Sensitivity analysis</i>	50
3.8	SOME EXAMPLES OF THE CONTROVERSIAL ASPECTS OF THE METHOD	51
3.9	RECENT METHODOLOGICAL DEVELOPMENTS (SINCE THE ORIGINAL REVIEW)	52
4	META-ANALYSIS FOR RARE DISEASE OUTCOMES: WHAT TO DO WHEN THERE ARE NO EXPOSED CASES?	56
4.1	INTRODUCTION	56
4.2	SUMMARY OF THE LITERATURE ON THE META-ANALYSIS OF SPARSE DATA SETS ...	56
4.2.1	<i>More recent methodological developments</i>	61
4.3	SYSTEMATIC PRESENTATION OF OPTIONS	63
4.3.1	<i>Analyses on the log scale</i>	64
4.3.2	<i>Analyses on the original scale.....</i>	67
4.3.3	<i>Transformations other than taking logarithms</i>	69

4.3.4	<i>Dealing with studies other than cohort studies</i>	70
4.3.5	<i>Meta-regression</i>	71
4.4	SIMULATIONS OF META-ANALYSES OF SMRs FROM COHORT STUDIES, WHEN EXPECTED NUMBER OF CASES IS SMALL.....	74
4.4.1	<i>Basic approach</i>	75
4.4.2	<i>Results when true summary SMR = 1.0</i>	75
4.4.3	<i>Results when true summary SMR = 1.5</i>	76
4.4.4	<i>Results when true summary SMR = 2.0</i>	77
4.4.5	<i>Indications for a continuity correction from the simulations</i>	78
4.5	SUMMARY OF APPROACHES	80
4.6	COMPARISON OF APPROACHES USING DATA FOR FORMALDEHYDE AND NASOPHARYNGEAL CANCER	81
4.7	COMPARISON OF APPROACHES USING DATA FOR FORMALDEHYDE AND SINONASAL CANCER.....	85
4.8	DISCUSSION AND CONCLUSIONS	87
5	REVIEW AND META-ANALYSIS OF THE EPIDEMIOLOGICAL EVIDENCE FOR THE CARCINOGENICITY OF FORMALDEHYDE IN RELATION TO LUNG CANCER, NASOPHARYNGEAL CANCER AND SINONASAL CANCER.....	89
5.1	INTRODUCTION	89
5.1.1	<i>Distribution of absorption of formaldehyde in the human body</i>	90
5.1.2	<i>The known and suspected risk factors for nasopharyngeal cancer, sinonasal cancer and lung cancer</i>	91
5.1.3	<i>Formaldehyde and its industrial use</i>	92
5.1.4	<i>Formaldehyde exposure in different occupational circumstances</i>	93
5.1.5	<i>Brief Summary of the toxicological evidence</i>	97
5.1.6	<i>IARC reviews of the carcinogenicity of formaldehyde</i>	99
5.1.7	<i>Previous meta-analyses of formaldehyde and respiratory cancer</i>	100
5.1.8	<i>Motivation and scope of this meta-analysis</i>	108
5.2	METHODS	108
5.2.1	<i>Study identification</i>	110
5.2.2	<i>Selection criteria for study inclusion in the meta-analysis</i>	111
5.2.3	<i>Extraction and coding of study characteristics</i>	112
5.2.4	<i>Meta-analytical methods</i>	114
5.3	RESULTS	115
5.3.1	<i>Characteristics of the studies analysed</i>	115
5.3.2	<i>Lung Cancer</i>	126
5.3.3	<i>Nasopharyngeal cancer</i>	141
5.3.4	<i>Sinonasal cancer</i>	145
5.4	DISCUSSION AND CONCLUSIONS	148
6	CONCLUSIONS.....	152
6.1	OVERVIEW.....	152
6.2	PAST PRACTICE OF META-ANALYSIS IN OCCUPATIONAL EPIDEMIOLOGY	153
6.3	REVIEW OF METHODS.....	153
6.4	RARE DISEASE OUTCOMES	154
6.5	META-ANALYSES OF FORMALDEHYDE	155
6.6	CLOSING REMARKS	158
	REFERENCES.....	160

APPENDIX 1 – SUMMARY OF THE EPIDEMIOLOGICAL STUDIES INCLUDED IN THE META-ANALYSIS IN CHAPTER 5.....	183
A1.1 THE NATIONAL CANCER INSTITUTE (NCI) COHORT	183
A1.2 THE NIOSH COHORT	185
A1.3 THE BRITISH COHORT	186
A1.4 OTHER INDUSTRIAL COHORT STUDIES.....	187
A1.5 COHORT STUDIES OF PATHOLOGISTS, ANATOMISTS, FUNERAL DIRECTORS AND EMBALMERS	189
A1.6 CASE-CONTROL STUDIES OF NASOPHARYNGEAL CANCER	190
A1.7 CASE-CONTROL STUDIES OF SINONASAL CANCER	193
A1.8 CASE-CONTROL STUDIES OF LUNG CANCER	195
A1.9 OTHER STUDIES INCLUDING REGISTRY-BASED	198

List of Tables and Figures

TABLE 4.1 - COMPARISON OF RESULTS WHEN TRUE SUMMARY SMR = 1.0	76
TABLE 4.2 - COMPARISON OF RESULTS WHEN TRUE SUMMARY SMR = 1.5	77
TABLE 4.3 - COMPARISON OF RESULTS WHEN TRUE SUMMARY SMR = 2	78
TABLE 4.4 - HEADLINE NASOPHARYNGEAL CANCER ANALYSES – ALL STUDIES	82
TABLE 4.5 - NASOPHARYNGEAL CANCER ANALYSES – COHORT STUDIES ONLY	84
TABLE 4.6 - HEADLINE SINONASAL CANCER ANALYSES – ALL STUDIES	85
TABLE 4.7 - SINONASAL CANCER ANALYSES – COHORT STUDIES ONLY.....	86
TABLE 5.1 - BRIEF OVERVIEW OF FORMALDEHYDE EXPOSURES IN WORKPLACES.....	94
TABLE 5.2 – RESULTS OF PREVIOUS META-ANALYSES OF FORMALDEHYDE IN RELATION TO NASOPHARYNGEAL CANCER, SINONASAL CANCER AND LUNG CANCER	101
TABLE 5.3 – MOOSE CHECKLIST FOR META-ANALYSES IN OBSERVATIONAL EPIDEMIOLOGY	109
FIGURE 5.1 - IDENTIFICATION OF STUDIES FOR INCLUDING THE META-ANALYSIS	116
TABLE 5.4 – OCCUPATIONAL COHORTS OF INDUSTRIAL WORKERS	117
TABLE 5.5 – OCCUPATIONAL COHORT STUDIES OF PATHOLOGISTS, ANATOMISTS, FUNERAL DIRECTORS AND EMBALMERS...	118
TABLE 5.6 – NASOPHARYNGEAL CANCER CASE-CONTROL STUDIES.....	119
TABLE 5.7 – SINONASAL CANCER CASE-CONTROL STUDIES.....	120
TABLE 5.8 – LUNG CANCER CASE-CONTROL STUDIES	122
TABLE 5.9 – OTHER STUDY TYPES	124
TABLE 5.10 – SUMMARY RISKS FOR META-ANALYSES OF ALL STUDIES OF LUNG CANCER AND SUBGROUPS – ANALYSES ON THE LOG SCALE	127

FIGURE 5.2 – RANDOM EFFECTS META-ANALYSIS OF LUNG CANCER FOR ALL RISK ESTIMATES.....	131
FIGURE 5.3 – RANDOM EFFECTS META-ANALYSIS OF LUNG CANCER BY YEAR OF PUBLICATION CATEGORY	132
FIGURE 5.4 - RANDOM EFFECTS META-ANALYSIS OF LUNG CANCER BY STUDY TYPE	133
FIGURE 5.5 - RANDOM EFFECTS META-ANALYSIS OF LUNG CANCER BY MORTALITY OR INCIDENCE STUDY.....	134
FIGURE 5.6 - RANDOM EFFECTS META-ANALYSIS OF LUNG CANCER BY STUDY LOCATION	135
FIGURE 5.7 - RANDOM EFFECTS META-ANALYSIS OF LUNG CANCER BY WHETHER OR NOT THE STUDY ADJUSTED FOR SMOKING	136
FIGURE 5.8 – RANDOM EFFECTS META-ANALYSIS OF LUNG CANCER BY STUDY QUALITY CATEGORY.....	137
TABLE 5.11 – META-RR FOR RANDOM EFFECTS MODEL EXCLUDING EACH STUDY IN TURN	138
FIGURE 5.9 – FUNNEL PLOT OF RELATIVE RISK ESTIMATES FOR LUNG CANCER IN RELATION TO OCCUPATIONAL EXPOSURE TO FORMALDEHYDE	139
FIGURE 5.10 - FUNNEL PLOT OF RELATIVE RISK ESTIMATES FOR LUNG CANCER IN RELATION TO OCCUPATIONAL EXPOSURE TO FORMALDEHYDE BY STUDY TYPE	140
FIGURE 5.11 – RANDOM EFFECTS META-ANALYSIS OF NASOPHARYNGEAL CANCER ON THE ORIGINAL SCALE	142
TABLE 5.12 – META-RR FOR FIXED EFFECT MODEL ON ORIGINAL SCALE EXCLUDING EACH STUDY IN TURN	143
FIGURE 5.12 – FUNNEL PLOT OF RELATIVE RISK ESTIMATES FOR NASOPHARYNGEAL CANCER IN RELATION TO OCCUPATIONAL EXPOSURE TO FORMALDEHYDE.....	145
FIGURE 5.13 – FIXED EFFECT META-ANALYSIS OF SINONASAL CANCER ON THE ORIGINAL SCALE	146
TABLE 5.14 – META-RR FOR FIXED EFFECT MODEL ON ORIGINAL SCALE EXCLUDING EACH STUDY IN TURN	147
FIGURE 5.14 – FUNNEL PLOT OF RELATIVE RISK ESTIMATES FOR SINONASAL CANCER IN RELATION TO OCCUPATIONAL EXPOSURE TO FORMALDEHYDE	148

1 INTRODUCTION

This chapter introduces the thesis, setting out the issues being addressed and the context in which they have been encountered.

1.1 THE META-ANALYSIS CONTROVERSY

The term meta-analysis was first coined in the mid-1970s [1]. Although the application of meta-analysis in the synthesis of the results of controlled clinical trials has become commonplace, its application in observational epidemiology remains controversial, because of the greater potential for biased effect estimates [2].

1.2 THE ENVIRONMENT OF OCCUPATIONAL REGULATORY RISK ASSESSMENT AND LIMIT SETTING

In occupational regulatory risk assessment in the United Kingdom, there remains an ongoing need to make decisions on the basis of published knowledge about whether a particular workplace hazard represents a risk to those exposed to it and at what level. Those decisions are often based on a qualitative assessment of the evidence. The first step is usually to determine whether or not the association between a workplace exposure and a particular disease outcome represents a causal association. If such a determination is made, a second step is to try to better quantify the nature of the exposure-response relationship. This usually involves characterising the risk at different levels of exposure. Of particular interest in the regulatory setting for determining appropriate occupational exposure limits, is the size or extent of the risk at the levels at which such exposure is encountered in the workplace.

1.3 THE AIMS OF THIS THESIS

The overall aim for this research is to identify the best approaches to the use of meta-analysis in the occupational health setting, and to summarise through meta-analysis the evidence as to whether or not occupational exposure to formaldehyde is an occupational carcinogen.

1.4 THE OUTLINE OF THIS THESIS

Chapter 2 contains a review of practice in occupational epidemiology and Chapter 3 contains a review of meta-analytical methods available for use in occupational epidemiology. Chapter 4 investigates the possible methodological approaches to meta-analysis that could be taken when there are risk estimates based on zero observed (observed) cases. Chapter 5 contains substantive meta-analyses of occupational exposure to formaldehyde, in relation to lung cancer, nasopharyngeal cancer and sinonasal cancer, making use of best practice in the application of meta-analysis methodology and applies methods that have been specially investigated for dealing with studies with zero (observed) cases. The final chapter draws the methodological and substantive findings together to form a view of the carcinogenicity of occupational exposure to formaldehyde in relation to respiratory cancers, discusses the strengths and limitations of the approach taken, and makes recommendations for future research.

2 LITERATURE REVIEW 1 – META-ANALYSIS: PAST PRACTICE IN OCCUPATIONAL EPIDEMIOLOGY

The main body of work contained in this chapter was published as a paper in the peer-reviewed scientific literature [3]. The completed cover sheet for the paper is in Section 2.1, and section 2.2 contains a brief update of practice since.

2.1 PUBLISHED REVIEW OF PAST PRACTICE IN OCCUPATIONAL EPIDEMIOLOGY

Cover sheet for each ‘research paper’ included in a research thesis

1. FOR A ‘RESEARCH PAPER’ ALREADY PUBLISHED

1.1. Where was the work published?

The peer-reviewed journal “Occupational Medicine”

1.2. When was the work published?

2004

1.2.1. *If the work was published prior to registration for your research degree, give a brief rationale for its inclusion*

Not applicable (registration was in 2000)

1.3. Was the work subject to academic peer review?

Yes

1.4. Have you retained the copyright for the work?

No, the work is subject to Crown copyright.

If yes, attach evidence of retention

Not applicable.

If no, or if the work is being included in its published format, attach evidence of permission from copyright holder (publisher or other author) to include work

This is not required, provided that I don’t state that HSE endorses the content of the publication (which I haven’t), I don’t use it to mislead (which I haven’t) and it doesn’t breach the Data Protection Act or the Privacy and Electronic Communications Regulations (which it doesn’t because it doesn’t use any personal data).

2. FOR A ‘RESEARCH PAPER’ PREPARED FOR PUBLICATION BUT NOT YET PUBLISHED

2.1. Where is the work intended to be published?

Not applicable

2.2. List the paper’s authors in the intended authorship order


Not applicable

2.3. Stage of publication – Not yet submitted/Submitted/ Undergoing revision from peer reviewers' comments/In press

Not applicable

3. For multi-authored work, give full details of your role in the research included in the paper and in the preparation of the paper. (Attach a further sheet if necessary)

The idea for undertaking this work, was principally mine, with input from my co-authors. I undertook the literature search and carried out the data analysis. Each author contributed to the discussion, with me leading and my co-authors ordered according to level of contribution to this aspect of the paper.



Candidate's signature

Supervisor or senior author's signature to confirm role as stated in (3):



Meta-analysis in occupational epidemiology: a review of practice

Damien M. McElvenny^{1,2}, Ben G. Armstrong¹, Lars Järup³ and Julian P. T. Higgins⁴

Objectives	To describe past practice in meta-analyses found in occupational epidemiology, identifying the major issues that should be considered by researchers planning a meta-analysis in this setting.
Methods	An electronic search of relevant online databases was undertaken. Papers were included in the review if they contained a statistical synthesis of risks in an occupational health setting.
Results	Sixty reports of meta-analyses were identified, mostly in cancer. The number of meta-analyses has increased consistently over the last 20 years. A majority of studies focused on a mean overall effect, although more than half of them also investigated heterogeneity of results. Both fixed effect and random effects meta-analysis models were employed, the former more often, and in eight studies used despite a statistically significant test for heterogeneity. A large proportion of the meta-analyses included different effect measures in the statistical synthesis, for example, including standardized mortality ratios (SMRs) and standardized incidence ratios. Most meta-analyses limited to a single type of effect measure focused on SMRs. The vast majority of meta-analyses combined all studies regardless of variation in the extent of information on exposures.
Conclusions	Meta-analyses in occupational epidemiology should properly explore and incorporate heterogeneity among studies. The meta-SMR is an important construct in this field, evidenced by a large proportion of cohort studies in the meta-analyses we identified. Controversy remains over the definition and validity of the meta-SMR. In addition, several other issues, notably dealing with heterogeneity in exposure, warrant further consideration.
Key words	Meta-analysis; occupational epidemiology; practice; review.
Received	7 July 2003
Revised	14 January 2004
Accepted	9 February 2004

Introduction

Meta-analyses are attractive to occupational health policy makers because they provide quantified statements in

relation to risks from occupational situations. Consequently, they are becoming more influential as an important foundation of occupational health policy. A meta-analysis is a review of previously conducted studies containing a statistical synthesis of individual study results in order to produce an overall average estimate of effect and/or to explore variation in estimates of effect. An introduction to the procedure is provided by Wong and Raabe [1], who consider its application to occupational epidemiology, and in particular to the meta-analysis of occupational cohort studies.

Although meta-analysis of controlled clinical trials is well established, for example, through initiatives such as

¹Public and Environmental Health Research Unit, London School of Hygiene and Tropical Medicine, London, UK.

²Epidemiology and Medical Statistics Unit, Health and Safety Executive, Merseyside, UK.

³Department of Epidemiology and Public Health, Imperial College Faculty of Medicine, London, UK.

⁴MRC Biostatistics Unit, Cambridge, UK.

Address for Correspondence: Damien McElvenny, Epidemiology and Medical Statistics Unit, Health & Safety Executive, Magdalen House, Stanley Precinct, Bootle, Merseyside L20 3QZ, UK. e-mail: damien.mcelvenny@hse.gsi.gov.uk

the Cochrane Collaboration [2], its application to observational studies has been controversial. Petitti [3] and Shapiro [4] discuss many of the important issues for meta-analysis of epidemiological studies in general. A major difference between the acceptability of meta-analyses of clinical trials and of observational epidemiological studies arises from use of randomization in the former in an attempt to eliminate bias and confounding. Shapiro points out that some meta-analysts have attempted to overcome the biases present in a meta-analysis of non-experimental studies by assigning quality weights to the component studies, in the hope that this can somehow overcome the biases. He concludes, however, that meta-analysis of non-experimental studies should be abandoned, because the information achieved by the meta-analytical approach cannot transcend the quality of the individual studies, and because meta-analysis is not being used in the way that its advocates recommend [4]. On the other hand, Petitti remarks that an inability to eliminate bias completely does not prevent epidemiologists from carrying out non-experimental studies, so that if meta-analyses of non-experimental data cannot overcome these problems, then all of non-experimental epidemiology should be abandoned. She further suggests that meta-analysis is an excellent tool for identifying possible reasons for variability and inconsistency between studies, even when it cannot resolve the inconsistency between them, and that formal evaluations of the quality and reproducibility of meta-analyses of non-experimental studies have not been done. Thus meta-analysis needs to be refined and its limitations defined rather than abandoned [3]. Simply estimating a mean effect size from a group of studies does not shed much light on a topic area and it has been suggested that meta-analysis should be treated as an aid to the qualitative review in critical comparison of different studies and for finding patterns among studies [1,5].

This paper aims to describe the most important features of meta-analyses in the occupational epidemiology field by collating those published up to and including 2001. A comprehensive search and examination of the studies has enabled us to identify the most important issues that should be considered by researchers planning a meta-analysis, and make some recommendations about the future use of meta-analysis in the occupational health setting.

Methods

An electronic search of online databases was undertaken. The databases searched were from 1975 to October 2001 were: NIOSH, HSELINE, MEDLINE OEM subset, NIOSHTIC (to September 1998 only), NIOSHTIC-2, Mental Health Abstracts, Applied Social Sciences Index

and Abstracts, Excerpta Medica, Medline, Psychinfo, Sociological Abstracts, Social Science Citation Index. Search terms for meta-analysis were meta*, meta-analysis and metaanalysis* and meta-analyt*. Work related studies were identified using search terms such as occupation*, work*, job*, industry*, labor* and labour*. Epidemiological studies were identified using free text term epidemiolog*. The searches were extended as far back as 1975, and restricted to the English language. A search was also carried out using the internet. This was supplemented by scanning the contents pages of the major occupational health and epidemiology journals, together with the citations in any of the meta-analyses found. Unpublished reports were not sought, and there remains the possibility that some relevant meta-analyses that have been completed are not included in this review, although we judge their likely impact on our conclusions to be minimal.

Once identified, papers were included if they contained a statistical synthesis of risks where the primary focus of attention was the occupational health setting. Pooled analyses, in which data on individual subjects from different studies are combined, were excluded. Pooled analyses offer a greater potential for analysis of dose-response relationships and person-specific risk factors than offered by meta-analysis of published reports [6], but such studies are very expensive and are much less common than meta-analyses of the literature.

For each of the identified meta-analyses included in this review, the following data were extracted:

- year of publication;
- nature of the disease or condition being investigated;
- occupational exposure or circumstances of interest;
- main aim of the meta-analysis;
- whether an overall estimate of effect was calculated;
- number and type of effect measures synthesized;
- type of statistical synthesis (whether fixed effect or random effects – see later for further details);
- statistical methods used in the synthesis;
- specific method of synthesis used for studies of the same type;
- whether the synthesis involved a weighted average, regression or another approach;
- whether different measures of effect were combined or analysed separately;
- type of heterogeneity test employed;
- result of the heterogeneity test;
- method employed for exploration of heterogeneity;
- characteristics examined in exploration of heterogeneity;
- basis of any quality weights used;
- method used for assessing publication bias;
- result of test for publication bias;
- method used for adjusting for publication bias;

- whether all studies were combined irrespective of exposure definitions in component studies;
- reasons for exclusions of any studies excluded on the basis of exposures;
- method of any dose-response analysis employed.

These data were then summarized in such a way as to inform the most important and specifically relevant meta-analysis issues in occupational epidemiology.

Results

Characteristics of the meta-analyses

Sixty meta-analyses were identified in the occupational epidemiological literature (Table 1). The number of published meta-analyses has increased steeply with successive quinquennial year groups since 1982 (Figure 1). The disease category most likely to be the subject of a

Table 1. Outcomes, main exposures and details of component studies of meta-analyses in occupational epidemiology published up to and including 2001

Ref.	First author	Year	Main outcome ^a	Main exposure ^b	No. and type of study included ^c	No. and type of risk estimates included ^d
[7]	Jackson	1985	PSYC	PSYC	56 NS	56 CCs
[8]	Morgan	1985	CAN	ASB	18 C	20 SMRs
[9]	Nelson	1986	CAN	ORG	19 NS	20 SMRs
[10]	Flowers	1988	INF	INF	6 CC	6 CIRR; 6 RR
[11]	Frumkin	1988	CAN	ASB	31 C	15 SMRs
[12]	Wong	1989	CAN	ORG	24 C	22 SMRs
[13]	Blair	1990	CAN	ORG	13 C 8 CC 2 NCC 9 RL	10 PMRs; 11 SMRs; 2 SIRs; 6 ORs
[14]	Frone	1990	PSYC	PSYC	7 CS	12 CCs
[15]	Johnson	1990	CAN	PEST	7 C 8 CC	7 SMRs; 7 ORs
[16]	Weiss	1990	CAN	ASB	20 C	Up to 12 SMRs/SIRs
[17]	Stock	1991	MSD	ERG	3 CC	3 ORs
[18]	Aldrich	1992	CAN	RAD	12 C 17 CC 8 RL	5 PMRs; 1 PIR; 6 ORs; 5 SMRs; 3 SIRs
[19]	Blair	1992	CAN	FARM	12 NS	24 risk estimates, mostly SMRs
[20]	Herbold	1993	CAN	ORG	6 C	6 SMRs
[21]	Parazzini	1993	REPR	RAD	6 CC 2 NCC	8 ORs
[22]	Partanen	1993	CAN	ORG	13 NS	10 PMRs; 11 SMRs; 2 SIRs; 6 ORs
[23]	Shore	1993	CAN	ORG	11 C	9 SMRs
[24]	Biggerstaff	1994	CAN	ETS	11 NS	Not clear – 8 O/E ratios
[25]	Homa	1994	CAN	ASB	20 C	20 SMRs
[26]	LeVois	1994	CAN	ETS	12 NS	Not clear – 15 O/E ratios
[27]	Partanen	1994	CAN	ORG	11 C 12 CC	Mixture of 16 ORs, SIRs, SMRs and O/E ratios
[28]	Rothman	1994	CAN	ORG	12 C	8 SMRs
[29]	Sjögren	1994	CAN	MET	2 C 2 CC 1 NCC	3 ORs; 2 SMRs
[30]	Fu	1995	CAN	MET	17 C 9 CC 4 NCC	15 O/E ratios from mixture of cohort and case-control studies
[31]	Henschler	1995	CAN	ORG	6 CC	6 ORs
[32]	Keller-Byrne	1995	CAN	FARM	6 C 10 CC 3 RL	6 SMRs; 10 ORs; 2 PMRs; 1 PCMR
[33]	Kheifets	1995	CAN	RAD	12 C 5 CC 1 NCC 11 RL	2 SMRs; 10 SIRs; 6 ORs; 5 PMR; 6 MORs
[34]	Chen	1996	VAR	ORG	55 C	46 SMRs; 8 RRs
[35]	Landtblom	1996	NSD	ORG	9 CC 1 RL 3 O	7 ORs; 1 PMR; 1 CC; 2 prevalence comparisons
[36]	Boivin	1997	REPR	NO	17 C 2 CC	17 RRs; 2 ORs
[37]	Collins	1997	CAN	ORG	11 C 22 CC 4 NCC 3 RL	11 SMRs; 3 PMRs; 26 ORs
[38]	Keller-Byrne	1997	CAN	FARM	8 C 13 CC 3 RL	19 studies, mainly cohort and case-control
[39]	Keller-Byrne	1997	CAN	FARM	5 CC 1 RL	1 PMR; 5 ORs
[40]	Kheifets	1997	CAN	RAD	24 C 14 CC 1 NCC 17 RL 1 CS 1 O	4 RRs; 16 ORs; 7 SMRs; 9 SIRs; 5 PMRs; 2 O/E ratios
[41]	Khuder	1997	CAN	FARM	8 C 19 CC 5 RL	5 PMRs; 19 ORs; 4 SIRs; 2 SMRs; 2 RRs
[42]	Lash	1997	CAN	ASB	15 C	23 SMRs
[43]	Moulin	1997	CAN	MET	20 C 16 CC	16 ORs; 20 SMRs
[44]	Paddle	1997	CAN	MET	12 C 2 RL	4 SMRs
[45]	Sparks	1997	VAR	WH	21 CS	19 CCs
[46]	Steenland	1997	VAR	SIL	12 C 1 CC 2 NCC 2 RL	11 SMRs; 6 ORs; 1 PMR; 1 PCMR; 1 SRR
[47]	Abbas	1998	MSD	ERG	12 CS 1 CC 4 O	11 cross-sectional studies; 4 surveys; 2 case-control studies
[48]	Acquavella	1998	CAN	FARM	16 C 9 CC 11 RL 1 O	12 SMRs; 11 PMRs; 8 ORs
[49]	Bhatia	1998	CAN	ORG	12 C 10 CC	10 ORs; 13 SMRs
[50]	Collins	1998	CAN	ORG	14 C 2 NCC 2 CC	4 ORs; 21 SMRs

Table 1. Continued

Ref.	First author	Year	Main outcome ^a	Main exposure ^b	No. and type of study included ^c	No. and type of risk estimates included ^d
[51]	McMartin	1998	REPR	ORG	6 C 4 CC	5 studies, mixture of cohort and case-control
[52]	Wells	1998	CAN	ETS	5 NS	18 studies, mixture of cohort and case-control
[53]	Goodman	1999	CAN	ASB	69 C	69 SMRs
[54]	Khuder	1999	CAN	FARM	17 NS	21 studies, mixture of cohort, case-control and other
[55]	Khuder	1999	CAN	FARM	7 C 13 CC 10 RL	30 studies, mixture of cohort, case-control and other
[56]	Lipsett	1999	CAN	ORG	18 C 19 CC 1 NCC 1 RL	39 studies, mixture of cohort, case-control and other
[57]	Stewart	1999	CAN	ORG	13 C 3 CC	9 SMRs; 3 ORs
[58]	Ballard	2000	CAN	AER	6 C	4 SMRs; 3 PMRs; 3 SIRs
[59]	Ojajarvi	2000	CAN	VAR	88 C 7 NCC 43 CC 23 RL	161 studies, mainly cohort and case-control
[60]	Wartenberg	2000	CAN	ORG	20 C	10 SMRs; 4 SIRs; 3 PMRs; 2 O/E ratios; 1 SMOR; 14 ORs; 11 other
[61]	Aryal	2001	NSD	ORG	1 C 7 CC	8 studies, all case-control except one cohort
[62]	Boffetta	2001	CAN	ORG	7 C 16 CC	30 studies, mixture of cohort, case-control and other
[63]	Collins	2001	CAN	ORG	8 C 2 CC 4 RL	8 SMRs; 3 PMRs; 1 PIR; 2 ORs
[64]	Greenberg	2001	VAR	VAR	181 C	461 SMRs
[65]	Neeraj	2001	RESP	ORG	2 C 18 CS 6 O	15 studies, 13 cross-sectional and two cohort
[66]	Ojajarvi	2001	CAN	ORG	24 C 7 CC 1 NCC 27 RL	19 SMRs; 7 SIRs; 18 PMRs; 4 PCMRs; 3 MORs; 4 ORs; 4 RRs

^aCAN = cancer; INF = infections; MSD = musculoskeletal disorders; NSD = nervous system diseases; PSYC = psychosocial disorders; REPR = reproductive effects; RESP = respiratory diseases; VAR = various diseases.

^bAVI = aviation; ASB = asbestos; ERG = ergonomic hazard; ETS = environmental tobacco smoke; FARM = farming; MET = metals or metallic compounds; ORG = organic compounds or solvents; NO = nitrous oxide; PEST = pesticides; PSYC = psychosocial hazard; RAD = radiation; SIL = silica; WH = working hours; VAR = various outcomes.

^cC = cohort study; CC = case-control (population or hospital-based) study; CS = cross-sectional study; NCC = nested case-control study; NS = not (clearly) stated; O = other study type; RL = record-linkage study.

^dNote details are included for the most inclusive meta-analysis presented in the paper. CC = correlation coefficients; SMR = standardized mortality ratio; CIRR = cumulative incidence rate ratio; RR = rate ratio; PMR = proportional mortality ratio; SIR = standardized incidence ratio; OR = odds ratio; PIR = proportionate incidence ratio; PCMR = proportionate cancer mortality ratio; MOR = mortality odds ratio; SRR = standardized rate ratio.

meta-analysis was cancer ($n = 46$), and the most common exposure or circumstances examined were organic solvents or compounds (23) and farming (8).

Not all of the meta-analyses were independent. For example, many of the meta-analyses that followed Blair *et al.* [19], which examined the risk of cancer from farming for many cancers, were assessments of the risks of specific types of cancer in relation to farming and included more recent studies. Also, the meta-analysis of formaldehyde and respiratory cancer by Partanen [22] is a re-analysis of the data included in the meta-analysis by Blair *et al.* [13].

Stated aims of the meta-analyses

Among the 34 meta-analyses with a clear and precisely stated aim, 30 stated an aim to estimate an overall measure of effect, 14 stated an aim of exploring heterogeneity and 10 reported aims as both of these. The proportion of studies that had exploration of heterogeneity as an aim has increased marginally with time (26% before 1996; 31% since 1997).

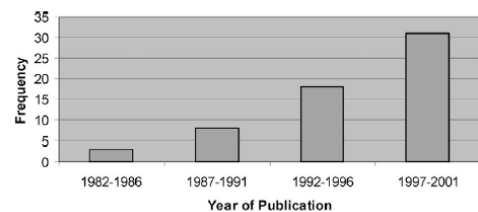


Figure 1. Frequency of meta-analyses in occupational epidemiology by year of publication.

Methods of statistical synthesis employed

A fixed effect analysis, which typically assumes each study is estimating the same underlying effect, was employed in 42 of the meta-analyses. Twenty of these also used a random effects approach to statistical synthesis, in which the studies are assumed to be estimating different underlying effects related across studies through some random distribution. Eight employed a random effects analysis only.

A large proportion (42/60) of the meta-analyses included different effect measures in the statistical synthesis. Of these, at least 27 contained either standardized mortality or incidence ratios (SMRs/SIRs). Of the 18 meta-analyses containing a single type of effect measure, the majority (13) synthesized SMRs via the meta-SMR, three pooled only odds ratios and two combined only average correlation coefficients (Table 1).

Testing and exploring heterogeneity

Of the 31 meta-analyses that stated which heterogeneity test they employed, most (23) employed the Cochran chi-squared statistic [67]. The only other method used more than twice was Breslow and Day's [68] method derived from logistic regression. Of the 20 meta-analyses employing only a fixed effect approach, eight reported a statistically significant heterogeneity test (Table 2).

Table 2. Summary of the methods for synthesis and investigation of heterogeneity for meta-analyses in occupational epidemiology published up to and including 2001

Ref.	First author	Method of synthesis (fixed effect or random effect/statistical method/regression or weighted average) ^a	Method of exploration of heterogeneity ^b	Result of heterogeneity test ^c
[7]	Jackson	FE/CC/WA	NONE	S & NS
[8]	Morgan	FE/MSMR/WA	NONE	NONE
[9]	Nelson	NS/NS/WA	SUBGRP	NONE
[10]	Flowers	FE/MH & ARD/WA	NONE	NS
[11]	Frumkin	RE/DL/WA	SUBGRP	SIG & NS
[12]	Wong	FE/MSMR/WA	NONE	NONE
[13]	Blair	FE/MRR/WA	SUBGRP	NONE
[14]	Frone	FE/CC/VC & WA	NONE	SIG & NS
[15]	Johnson	FE/CPMR & MH/WA	NONE	S
[16]	Weiss	FE/MSMR/WA	SUBGRP	NONE
[17]	Stock	NS/NS/NS	NONE	NONE
[18]	Aldrich	NS/NS/NS	NONE	NONE
[19]	Blair	FE/MRR & ST/WA	NONE	NONE
[20]	Herbold	FE & RE/PV & MSMR/WA	NONE	NONE
[21]	Parazzini	FE/MH/WA	NONE	NONE
[22]	Partanen	FE/MRR/WA	SUBGRP	NS
[23]	Shore	FE & RE/MSMR/WA	SUBGRP	SIG & NS
[24]	Biggerstaff	FE & RE/MH & C & DM/WA & R & B	REGR	SIG & NS
[25]	Homa	FE & RE/MSMR & MR/WA & R	REGR	SIG & NS
[26]	LeVois	NS/MRR/WA	SUBGRP	NONE
[27]	Partanen	NS/NS/WA	SUBGRP	SIG & NS
[28]	Rothman	NS/MSMR/WA	NONE	NONE
[29]	Sjögren	NS/NS/NS	NONE	NONE
[30]	Fu	FE & RE/DL & MRR/WA	SUBGRP	SIG & NS
[31]	Henschler	NS/NS/WA	NONE	NONE
[32]	Keller-Byrne	FE & RE/DL/WA	SUBGRP	SIG & NS
[33]	Kheifets	FE & RE/DL & MR/WA & R	REGR & SUBGRP	SIG & NS
[34]	Chen	FE & RE/MSMR/WA	NONE	SIG & NS
[35]	Landtblom	FE/MRR/WA	SUBGRP	NONE
[36]	Boivin	FE/NS/WA	SUBGRP	NONE
[37]	Collins	FE/MRR/WA	SUBGRP	SIG & NS
[38]	Keller-Byrne	FE & RE/DL/WA	NONE	SIG
[39]	Keller-Byrne	RE/DL/WA	NONE	SIG
[40]	Kheifets	FE & RE/DL & MR/WA & R	REGR & SUBGRP	SIG
[41]	Khuder	RE/DL/WA	SUBGRP	SIG
[42]	Lash	FE & RE/ML/WA	REGR	SIG
[43]	Moulin	FE/MRR/WA	SUBGRP	NS
[44]	Paddle	NS/NS/WA	NONE	NONE
[45]	Sparks	NS/CC/WA	NONE	NONE
[46]	Steenland	RE/DL/WA	NONE	NONE
[47]	Abbas	FE & RE/DL & MR/WA & R	REGR	NONE
[48]	Acquavella	FE & RE/MRR/WA & R	SUBGRP	SIG & NS
[49]	Bhatia	FE & RE/MRR/WA	SUBGRP	SIG & NS
[50]	Collins	FE & RE/MRR/WA	SUBGRP	S & NS
[51]	McMartin	FE/MH/WA	NONE	NS
[52]	Wells	FE/NS/WA	NONE	NONE
[53]	Goodman	FE/MSMR/WA	SUBGRP	SIG
[54]	Khuder	RE/DL/WA	SUBGRP	SIG

Table 2. Continued

Ref.	First author	Method of synthesis (fixed effect or random effect/statistical method/regression or weighted average) ^a	Method of exploration of heterogeneity ^b	Result of heterogeneity test ^c
[55]	Khuder	RE/DL/WA	SUBGRP	SIG
[56]	Lipsett	RE/DL/WA	REGR & SUBGRP	SIG
[57]	Stewart	FE/W/WA	NONE	SIG & NS
[58]	Ballard	FE & RE/DL/WA	SUBGRP	SIG & NS
[59]	Ojajarvi	FE & RE/DL/WA	SUBGRP	SIG & NS
[60]	Wartenberg	FE2/MRR/WA	NONE	NONE
[61]	Aryal	RE/DL/WA	NONE	SIG
[62]	Boffetta	FE/MR/R	REGR & SUBGRP	SIG & NS
[63]	Collins	FE & RE/MRR & WB/WA	SUBGRP	SIG & NS
[64]	Greenberg	FE & RE/MSMR/WA	SUBGRP	SIG & NS
[65]	Neeraj	FE/BD & AB/WA	MU	SIG & NS
[66]	Ojajarvi	FE & RE/MR & DL/R & WA	REGR	SIG & NS

^aFE = fixed effect; NS = not (clearly) stated; RE = random effects; AB = Abramson; ARD = adjusted rate differences; BD = Breslow and Day; C = Carlin (Bayesian); CC = combined correlation coefficients; CPMR = crude proportional mortality ratio; DL = DerSimonian and Laird; DM = DuMouchel (Bayesian); MH = Mantel-Haenszel; ML = maximum likelihood; MRR = meta-relative risk; MSMR = meta-standardized mortality ratio; PV = combination of *P*-values; W = Woolf; WB = Wang and Bushman; B = Bayesian approach; R = regression; VC = vote counting; WA = weighted average.

^bNONE = none presented; REGR = regression analysis; SUBGRP = subgroup analysis.

^cNONE = not done or not presented; SIG = statistically significant; NS = not significant.

Just over half (34) of the meta-analyses presented an exploration of heterogeneity, 20 via an informal subgroup analysis, five by regression analysis and four via both subgroup and regression analysis (Table 2). In addition to study design, the characteristics examined included geographical region, industry, and various possible confounding factors.

Publication and reporting bias

Evidence of or against publication bias was sought in 17 meta-analyses, with only one resulting in a determination that there was significant publication bias present, one producing a mixture of significant and non-significant results, and 10 producing a non-significant result. In six further studies, the presence of publication bias was examined informally with two of these finding some evidence for its existence. A variety of methods were employed, including funnel plots [69]. None of the meta-analyses formally adjusted for publication bias.

Exposure data

The vast majority of meta-analyses (50) combined all studies regardless of the extent of information supplied on occupational exposures of the study subjects. Thirteen of these 50 meta-analyses considered effect estimates in groups of studies according to level of exposure, to look for a dose response relationship. Only two meta-analyses examined this formally.

Discussion

Our review of published meta-analyses in the occu-

pational epidemiology field identified 60 meta-analyses, and allows us to describe and critique current practice. We might do this by taking as a basis the view of Greenland [5], with which we agree, that meta-analysis should be treated as an aid to comparisons among studies, rather than focused solely on estimation of an overall mean. It seems that meta-analysis in occupational health has moved somewhat towards this viewpoint, as evidenced by the increasing use of meta-regression and examination of subgroups of studies. Nevertheless, focus on overall mean effects is still prevalent, and the practice of employing a fixed effect analysis in the presence of significant heterogeneity is a cause for concern. The stated rationale for employing a random effects approach seems mainly to centre on producing wider confidence intervals around overall risk estimates. This is desirable, but does not obviate the need to investigate reasons for inconsistent results.

Several issues have emerged from this review as key ones for occupational meta-analysis. Most prominent was that of variations in definitions of exposure across studies and lack of information on exposure, which make comparisons difficult. Apart from those meta-analyses specifically focused on deriving a dose-response, the vast majority of the remaining meta-analyses combined study results where the exposure measures may well not be comparable: for example, studies with semi-quantitative exposure estimates being combined with studies where exposures were simply defined by job title.

Other aspects of the statistical synthesis were recognized as having the potential to introduce bias. Examples include the combination of risk estimates adjusted with those unadjusted for confounders (as well

as the control for different sets of confounders) and those restricted with those unrestricted for latency (time from first exposure to onset of the disease of interest). Some of the biases were addressed by the use of meta-regression based on study characteristics. However, a cautionary note has been reaffirmed in the examination of heterogeneity between studies, since there will often be insufficient data for this to be investigated reliably using statistical methods [70].

A variety of approaches were employed in estimating the overall measure of effect. Apart from the generic approaches, in which all effect measures were assumed to approximate a relative risk, the most important for occupational epidemiology is the use of the meta-SMR for combining the results of cohort studies. Occupational epidemiology is unusual in the preponderance of the SMR as an effect measure. The use of SMRs in meta-analyses raises two questions. First, is it valid to compare and combine SMRs? Secondly, if SMRs are to be combined in a meta-analysis, how should this be done? It has been argued that the SMR as an effect estimate is inherently non-comparable across individual studies [71]. Others are less critical of comparing SMRs, noting that SMRs can be validly compared where the underlying stratum-specific death rates for each exposure class (or study) are in constant proportion to those of external standard rates, although these authors acknowledge the possibility of bias if this condition is not met [72,73]. Others have considered that large differences in the age or sex structure of populations—the other condition for bias in comparing SMRs—in practice rarely occur to a substantial degree [74]. Few of the meta-analyses in our review considered the validity of comparing and combining SMRs. The non-comparability of SMRs may to an extent be encompassed in a random effects meta-analysis model [11]. This allows the studies to be estimating different, yet related, underlying effects. It typically yields an estimate of the average of these effects, accompanied by a suitably wider confidence interval than the fixed effect approach [11]. Fixed effect meta-analyses continue to be undertaken routinely, either in the absence of a test for heterogeneity or, more alarmingly, when there is evidence for heterogeneity. Furthermore, the Cochran chi-squared test for heterogeneity is not very powerful [75], yet it continues to be used as a justification for adopting only a fixed effect approach to the analysis.

Publication bias was not much considered in these studies. Despite this, many of the authors employing funnel plots did not acknowledge that asymmetry in funnel plots can be due to an exaggeration of effect size in small studies of low quality, as well as publication bias [76]. It may be, however, that reporting bias is more of an issue for occupational epidemiology, given the lack of guidelines for the reporting of epidemiological studies carried out in the occupational field [44].

It was noticeable from this review that there were major variations in the level and type of information reported in the meta-analyses. There was little indication that reporting proposals [77] are enforced by reviewers and journal editors. Also clear is that many of the meta-analyses carried out would not meet the criteria required according to the guidelines set out by Blair *et al.* [78] for meta-analyses in environmental epidemiology.

Conclusions

Meta-analyses are becoming increasingly popular in occupational epidemiology. However, Petitti's note that meta-analysis in this field needs to be refined and its limitations defined appears not to have been heeded. Future meta-analyses in occupational epidemiology should ensure that heterogeneity of study results is properly incorporated or explored, that variations in the quality of studies is acknowledged, and that the possible impact of reporting is investigated and reported.

Some issues specifically relevant to occupational epidemiology require further evaluation. The two principle issues are (i) the integration of diverse measures of occupational exposure, and important sources of variation across occupational epidemiological studies; and (ii) the comparability and combination of SMRs, where further investigation is required to determine when a meta-SMR might be appropriate as an overall estimate of effect.

Acknowledgements

We would like to thank JT Hodgson for commenting on the various drafts of this paper. We would also like to thank the Search Team from the Health & Safety Executive's Sheffield Information Centre for carrying out the literature searches, and Bernie Good, Tracy Hamilton, Jan Harding, Karen Hughes, Phil Irwin and Lori Woods for administrative support.

References

1. Wong O, Raabe GK. Application of meta-analysis in reviewing occupational cohort studies. *Occup Environ Med* 1996;53:793–800.
2. Chalmers I. The Cochrane Collaboration: preparing, maintaining and disseminating systematic reviews of the effects of health care. *Ann NY Acad Sci* 1993;703:156–163.
3. Petitti DB. Of babies and bathwater. *Am J Epidemiol* 1994;9:779–782.
4. Shapiro S. Meta-analysis/Shmeta-analysis. *Am J Epidemiol* 1994;9:771–778.
5. Greenland S. Can meta-analysis be salvaged? *Am J Epidemiol* 1994;9:783–787.
6. Checkoway H. Data pooling in occupational studies. *J Occup Med* 1991;33:1257–1260.

7. Jackson SE, Schuler RS. A meta-analysis and conceptual critique of research on role ambiguity and role conflict in work settings. *Organiz Behav Hum Decision Processes* 1985;36:16–78.
8. Morgan RW, Foliart DE, Wong O. Asbestos and gastrointestinal cancer: a review of the literature. *West J Med* 1985;143:60–65.
9. Nelson N, Levine RJ, Albert RA, et al. Contribution of formaldehyde to respiratory cancer. *Environ Health Perspectives* 1986;70:23–35.
10. Flowers RH III, Torner JC, Farr BM. Primary cytomegalovirus infection in pediatric nurses: a meta-analysis. *Control Hosp Epidemiol* 1988;9:491–496.
11. Frumkin H, Berlin J. Asbestos exposure and gastrointestinal malignancy: review and meta-analysis. *Am J Ind Med* 1988;14:79–95.
12. Wong O, Raabe GK. Critical review of cancer epidemiology in petroleum industry employees, with a quantitative meta-analysis by cancer site. *Am J Ind Med* 1989;15:283–310.
13. Blair A, Saracci R, Stewart PA, Hayes RB, Shy C. Epidemiologic evidence on the relationship between formaldehyde exposure and cancer. *Scand J Work Environ Health* 1990;16:381–393.
14. Frone MR. Intolerance of ambiguity as a moderator of the occupational role stress-strain relationship: a meta-analysis. *J Organiz Behav* 1990;11:309–320.
15. Johnson CC, Feingold M, Tilley B. A meta-analysis of exposure to phenoxy acid herbicides and chlorophenols in relation to soft tissue sarcoma. *Int Arch Occup Environ Health* 1990;62:513–520.
16. Weiss W. Asbestos and colorectal cancer. *Gastroenterology* 1990;99:876–884.
17. Stock S. Workplace ergonomic factors and the development of musculoskeletal disorders of the neck and upper limbs: a meta-analysis. *Am J Ind Med* 1991;19:87–107.
18. Aldrich TE, Laborde D, Griffith J, Easterly C. A meta-analysis of the epidemiological evidence regarding human health risk associated with exposure to electromagnetic fields. *Electro Magnetobiol* 1992;11:127–143.
19. Blair A, Zahm SH, Pearce NE, Heineman EF, Fraumeni JF Jr. Clues to cancer etiology from studies of farmers. *Scand J Work Environ Health* 1992;18:209–215.
20. Herbold, M. Meta-analysis of environmental and occupational epidemiological studies: a method demonstrated using the carcinogenicity of PCBs as an example. *Soz Präventivmed* 1993;38:185–189.
21. Parazzini F, Luchini L, La Vecchia C, Crosignani PG. Video display terminal use during pregnancy and reproductive outcome—a meta-analysis. *J Epidemiol Community Health* 1993;47:265–268.
22. Partanen T. Formaldehyde exposure and respiratory cancer—a meta-analysis of the epidemiologic evidence. *Scand J Work Environ Health* 1993;19:8–15.
23. Shore RE, Gardner MJ, Pannett B. Ethylene oxide: an assessment of the epidemiological evidence on carcinogenicity. *Br J Ind Med* 1993;50:971–977.
24. Biggerstaff BJ, Tweedie RL, Mengersen KL. Passive smoking in the workplace. *Int Arch Occup Environ Health* 1994;66:269–277.
25. Homa DM, Garabrant DH, Gillespie BW. A meta-analysis of colorectal cancer and asbestos exposure. *Am J Epidemiol* 1994;139:1210–1222.
26. LeVois M, Layard MW. Inconsistency between workplace and spousal studies of environmental tobacco smoke and lung cancer. *Regul Toxicol Pharmacol* 1994;19:309–316.
27. Partanen T, Boffetta P. Cancer risk in asphalt workers and roofers: review and meta-analysis of epidemiologic studies. *Am J Ind Med* 1994;26:721–740.
28. Rothman KJ. Cancer occurrence among workers exposed to acrylonitrile. *Scand J Work Environ Health* 1994;20:313–321.
29. Sjögren B, Hansen KS, Kjuus H, Persson P-G. Exposure to stainless steel welding fumes and lung cancer: a meta-analysis. *Occup Environ Med* 1994;51:335–336.
30. Fu H, Boffetta P. Cancer and occupational exposure to inorganic lead compounds: a meta-analysis of published data. *Occup Environ Med* 1995;52:73–81.
31. Henschler D, Vamvakas S, Lammert M, et al. Increased incidence of renal cell tumours in a cohort of cardboard workers exposed to trichloroethylene. *Arch Toxicol* 1995;70:131–133.
32. Keller-Byrne JE, Khuder SA, Schaub EA. Meta-analysis of leukaemia and farming. *Environ Res* 1995;71:1–10.
33. Kheifets LI, Afifi AA, Buffler PA, Zhang ZW. Occupational electric and magnetic field exposure and brain cancer: a meta-analysis. *J Occup Environ Med* 1995;37:1327–1341.
34. Chen R, Seaton A. A meta-analysis of mortality among workers exposed to organic solvents. *Occup Med* 1996;46:337–344.
35. Landtblom A-M, Flodin U, Soderfeldt B, Wolfson C, Axelsson O. Organic solvents and multiple sclerosis: a synthesis of the current evidence. *Epidemiology* 1996;7:429–433.
36. Boivin J-F. Risk of spontaneous abortion in women occupationally exposed to anaesthetic gases: a meta-analysis. *Occup Environ Med* 1997;54:541–548.
37. Collins JJ, Acquavella JF, Esmen NA. An updated meta-analysis of formaldehyde exposure and upper respiratory tract cancers. *J Occup Environ Med* 1997;39:639–651.
38. Keller-Byrne EA, Khuder SA, Schaub EA. Meta-analyses of prostate cancer and farming. *Am J Ind Med* 1997;31:580–586.
39. Keller-Byrne JE, Khuder SA, Schaub EA, McAfee O. A meta-analysis of non-Hodgkin's lymphoma among farmers in the central United States. *Am J Ind Med* 1997;31:442–444.
40. Kheifets LI, Afifi AA, Buffler PA, Zhang ZW, Matkin CC. Occupational electric and magnetic field exposure and leukaemia. *J Occup Environ Med* 1997;39:1074–1091.
41. Khuder SA, Mutgi AB. Meta-analyses of multiple myeloma and farming. *Am J Ind Med* 1997;32:510–516.
42. Lash TL, Crouch EAC, Green LC. A meta-analysis of the relation between cumulative exposure to asbestos and relative risk of lung cancer. *Occup Environ Med* 1997;54:254–263.
43. Moulin JJ. A meta-analysis of epidemiologic studies of lung cancer in welders. *Scand J Work Environ Health* 1997;23:104–113.
44. Paddle GM. Metaanalysis as an epidemiological tool and

- its application to studies of chromium. *Regulatory Toxicol Pharmacol* 1997;26:S42–S50.
45. Sparks K, Cooper C, Fried Y, Shirom A. The effects of hours of work on health. *J Occup Organizational Psychol* 1997;70:391–408.
 46. Steenland K, Stayner L. Silica, asbestos, man-made mineral fibers and cancer. *Cancer Causes Control* 1997;8:491–503.
 47. Abbas MAF, Afifi AA, Zhang ZW, Kraus JF. Meta-analysis of published studies of carpal tunnel syndrome. *Int J Occup Environ Health* 1998;4:160–167.
 48. Acquavella J, Olsen G, Cole P, et al. Cancer among farmers: a meta-analysis. *Ann Epidemiol* 1998;8:64–74.
 49. Bhatia R, Lopipero P, Smith AH. Diesel exhaust exposure and lung cancer. *Epidemiology* 1998;9:84–91.
 50. Collins JJ, Acquavella JF. Review and meta-analysis of studies of acrylonitrile workers. *Scand J Work Environ Health* 1998;24(Suppl. 2):71–80.
 51. McMartin KI, Chu M, Kopecky E, Einarson TR, Koren G. Pregnancy outcome following maternal organic solvent exposure: a meta-analysis of epidemiologic studies. *Am J Ind Med* 1998;34:288–292.
 52. Wells AJ. Lung cancer from passive smoking at work. *Am J Public Health* 1988;88:1025–1029.
 53. Goodman M, Morgan RW, Ray R, Malloy CD, Zhao K. Cancer in asbestos-exposed occupational cohorts: a meta-analysis. *Cancer Causes Control* 1999;10:453–465.
 54. Khuder SA. Etiologic clues to lip cancer from epidemiologic studies on farmers. *Scand J Work Environ Health* 1999;25:125–130.
 55. Khuder SA, Mutgi AB, Schaub EA, Tano BDK. Meta-analysis of Hodgkin's disease among farmers. *Scand J Work Environ Health* 1999;25:436–441.
 56. Lipsett M, Campleman S. Occupational exposure to diesel exhaust and lung cancer: a meta-analysis. *Am J Publ Health* 1999;89:1009–1017.
 57. Stewart RE, Dennis LK, Dawson DV, Resnick MI. A meta-analysis of risk estimates for prostate cancer related to tire and rubber manufacturing operations. *J Occup Environ Med* 1999;41:1079–1084.
 58. Ballard T, Lagorio S, De Angelis G, Verdecchia A. Cancer incidence and mortality among flight personnel: a meta-analysis. *Aviat Space Environ Med* 2000;71:216–224.
 59. Ojajarvi IA, Partanen TJ, Ahlbom A, et al. Occupational exposures and pancreatic cancer: a meta-analysis. *Occup Environ Med* 2000;57:316–324.
 60. Wartenberg D, Reyner D, Scott CS. Trichloroethylene and cancer: epidemiologic evidence. *Environ Health Perspect* 2000;108(Suppl. 2):161–176.
 61. Aryal BK, Khuder SA, Schaub EA. Meta-analysis of systemic sclerosis and exposure to solvents. *Am J Ind Med* 2001;40:271–274.
 62. Boffetta P, Silverman DT. A meta-analysis of bladder cancer and diesel exhaust exposure. *Epidemiology* 2001;12:125–130.
 63. Collins JJ, Esman NA, Hall TA. A review and meta-analysis of formaldehyde exposure and pancreatic cancer. *Am J Ind Med* 2001;39:336–345.
 64. Greenberg RS, Mandel JS, Pastides H, Britton NL, Rudenko L, Starr TB. A meta-analysis of cohort studies describing mortality and cancer incidence among chemical workers in the United States and Western Europe. *Epidemiology* 2001;12:727–740.
 65. Neeraj M, Rastogi SK, Srivastava AK. Occupational epidemiological studies of respiratory disorders due to exposure to organic solvents—systematic review with meta-analysis. *Ind J Occup Environ Med* 2001;5:118–126.
 66. Ojajarvi A, Partanen T, Ahlbom A, et al. Risk of pancreatic cancer in workers exposed to chlorinated hydrocarbons. *Am J Epidemiol* 2001;153:841–850.
 67. Cochran WG. The combination of estimates from different experiments. *Biometrics* 1954;10:101–129.
 68. Breslow NE, Day NE. *Statistical Methods in Cancer Research. Vol. II. The Design and Analysis of Cohort Studies*. New York: Oxford University Press, 1987; 406.
 69. Egger M, Smith GD, Schneider M, Minder C. Bias in meta-analysis detected by a simple, graphical test. *Br Med J* 1997;315:629–634.
 70. Higgins J, Thompson S, Deeks J, Altman D. Statistical heterogeneity in systematic reviews of clinical trials: a critical appraisal of guidelines and practice. *J Health Serv Res Policy* 2002;7:51–61.
 71. Myers JE and Thompson ML. Meta-analysis and occupational epidemiology. *Occup Med* 1998;48:99–101.
 72. Breslow NE and Day NE. *Statistical Methods in Cancer Research. Vol. II. The Design and Analysis of Cohort Studies*. Oxford: Oxford University Press, 1987; 151–152.
 73. Armstrong BG. Comparing standardized mortality ratios. *Ann Epidemiol* 1995;5:60–64.
 74. Järup L, Best N. Editorial comments on geographical differences in cancer incidence in the Belgian province of Limburg by Bruntinx and colleagues. *Eur J Cancer* 2003;39:1973–1975.
 75. Higgins JPT, Thompson SG, Deeks JJ, Altman DG. Measuring inconsistency in meta-analysis. *Br Med J* 2003;327:557–560.
 76. Sterne JAC, Eggar M, Smith GD. Investigating and dealing with publication and other biases. In: Eggar M, Smith GD, Altman DG, eds. *Systematic Reviews in Health Care: Meta-analysis in Context*. London: BMJ Books, 2001; 189–208.
 77. Stroup DF, Berlin JA, Morton SC, et al. Meta-analysis of observational studies in epidemiology: a proposal for reporting. Meta-analysis in observational studies (MOOSE) group. *J Am Med Assoc* 2000;283:2008–2012.
 78. Blair A, Burg J, Foran J, et al. Guidelines for application of meta-analysis in environmental epidemiology. *Regul Toxicol Pharmacol* 1995;22:189–197.

2.2 PAST PRACTICE IN OCCUPATIONAL EPIDEMIOLOGY: A BRIEF UPDATE

A recently as 2013, it has been restated that there is a long history of debate about the value of meta-analysis for occupational cohort studies [4], reminding us of Shapiro's argument that "meta-analysis of published non-experimental data should be abandoned", in which he reasoned that relative risks of low magnitude, say less than 2, are virtually beyond the resolving power of the epidemiological microscope because we can seldom eliminate all sources of bias. Because the pooling of studies in a meta-

analysis increases statistical power, the pooled estimate can easily become significant and thus incorrectly taken as an indication of causality even though the biases in the included studies may not have been taken into account [5]. Others have argued that the method of meta-analysis is important but should be applied appropriately taking into account the biases in the individual studies [6]. The more methodological research on meta-analysis in occupational epidemiological which was called for around 10 years ago [3] has yet to be carried out, despite that huge increase in applications of the method that have occurred during that time. In particular, the meta-SMR remains a complicated concept, for example, because of its incorporation of standardization on completely different and often non-comparable populations [4].

3 LITERATURE REVIEW 2 - META-ANALYSIS METHODS AVAILABLE FOR THE OCCUPATIONAL HEALTH SETTING

3.1 INTRODUCTION: SCOPE AND STRUCTURE OF THE REVIEW

The original purpose of this chapter is to describe published statistical methods for meta-analysis, emphasising those that are specifically relevant or important to the occupational health setting, or at least those that were when the search was done in 2001. This has been supplemented with a brief summary of some of the key methodological approaches that have been developed since the original review.

After a description of the search methodology employed to identify potentially relevant references, this Chapter is split into methods that attempt to estimate an overall mean effect measure (fixed effect and random effects) and those that assess and explore between study heterogeneity. Finally, the Chapter sets out some methods for the identification and adjustment for the effects of publication bias and other relevant methods.

3.2 SEARCH METHODOLOGY

The following on-line databases were searched from the earliest year available to Oct 2001: OSHROM (NIOSH, HSELINE, MEDLINE OEM subset, NIOSHTIC (to Sep 1998 only), NIOSHTIC-2); DIALOG (Mental Health Abstracts); DATASTAR (Applied Social Sciences Index and Abstracts, Excerpta Medica, Medline, Psychinfo, Sociological Abstracts, Social Science Citation Index) using the free text terms meta-analys*, meta-analyt*, metaanalys* and metaanalyt* combined with method*. In addition, methodological citations from the meta-analysis examples identified in

Chapter 2 and papers already held and their citations, together with the contents pages of Statistics in Medicine and Biometrics were used.

3.3 METHODS FOR ESTIMATING MEAN MEASURE OF EFFECT

Historically meta-analyses, particularly in the occupational health field, initially concentrated on estimating an overall average effect estimate across all the studies relevant to the hypothesis under question. This Section sets out the main statistical measures of mean effect of relevance for the occupational health setting.

3.3.1 Fixed effect methods

These methods have an underlying assumption that each of the studies included in the meta-analysis is estimating a single true underlying “fixed” measure of effect. This can be expressed mathematically as

$$Y_i = \mu + e_i$$

where Y_i is the observed effect from study i , μ is the population effect size and e_i is the sampling error.

3.3.1.1 Combining generic effect measures by means of a weighted average approach

If Y denotes the generic effect measure and w the reciprocal of its variance, then an estimator of an assumed common underlying effect size, for studies $i = 1, \dots, k$, is

$$\bar{Y} = \frac{\sum_{i=1}^k w_i Y_i}{\sum_{i=1}^k w_i},$$

where

$$w_i = \frac{1}{v_i} = \frac{1}{\text{var}(Y_i)},$$

and an associated approximate $100(1-\alpha)\%$ confidence interval is given by

$$\bar{Y} \pm \frac{z_{1-\alpha/2}}{\sqrt{\sum_{i=1}^k w_i}} \quad [7].$$

If Y_i is assumed to be a ratio (a relative risk), then by putting

$$Y_i = \log RR_i$$

the mean relative risk on a log scale can be expressed as

$$\overline{\log RR} = \frac{\sum_{i=1}^k w_i \log RR_i}{\sum_{i=1}^k w_i},$$

where

$$w_i = \frac{1}{\text{var}(\log RR_i)}.$$

In practice, because variances or standard errors of risk estimates may not be presented in scientific papers, an estimate of the variance of $\log RR_i$ can be calculated from the associated $100(1-\alpha)\%$ confidence interval using

$$\text{var}(\log RR_i) = \left[\frac{\log RR_{upper} - \log RR_{lower}}{2z_{1-\alpha/2}} \right]^2.$$

The variance of the combined logRR or meta-logRR is

$$\text{var}(\overline{\log RR}) = \frac{1}{\sum_{i=1}^k w_i}.$$

Thus, assuming normality of RR on the log scale, a $100(1-\alpha)\%$ confidence interval can be expressed as

$$\overline{\log RR} \pm z_{1-\alpha/2} \sqrt{\text{var}(\overline{\log RR})} \quad [8].$$

3.3.1.2 Combining Standardised Mortality Ratios (SMRs)

A SMR is a measure of RR that is particularly important in the occupational health setting. It is aggregated in those meta-analyses that are the combination of the results of cohort studies where the analyses are based on comparison with an external population.

An overall SMR can be calculated, using

$$Y_i = \log(SMR_i),$$

from

$$\overline{\log SMR} = \frac{\sum_{i=1}^k w_i \log(SMR_i)}{\sum_{i=1}^k w_i},$$

as

$$meta-SMR_1 = \exp(\overline{\log SMR}),$$

with

$$w_i = \frac{1}{\text{var}(\log SMR_i)}$$

and

$$\text{var}(\log SMR_i) = \frac{1}{O_i} \quad [9].$$

An approximate $100(1-\alpha)\%$ confidence interval for the $\overline{\log SMR}$ is given by

$$\overline{\log SMR} \pm \frac{z_{1-\alpha/2}}{\sqrt{\sum_{i=1}^k O_i}},$$

and thus for the meta-SMR₁ is

$$meta - SMR_1 = \exp \left(\overline{\log SMR} \pm \frac{z_{1-\alpha/2}}{\sqrt{\sum_{i=1}^k O_i}} \right) \quad [9].$$

Note that if the number of observed and expected events is available from the individual studies, then an alternative form of the meta-SMR can also be calculated simply by using

$$meta - SMR_2 = \frac{\sum_{i=1}^k O_i}{\sum_{i=1}^k E_i} \quad [10],$$

which is algebraically equivalent to

$$meta - SMR_2 = \frac{\sum_{i=1}^k w_i SMR_i}{\sum_{i=1}^k w_i},$$

where

$$w_i = \frac{E_i}{\sum_{i=1}^k E_i}.$$

Note that meta-SMR₁ and meta-SMR₂ are algebraically different and meta-SMR₂ is to be preferred when the data are available to calculate it.

3.3.1.3 Combining Odds Ratios (ORs)

These methods are used when the studies to be combined are all case-control studies, and have the advantage of being able to deal with ORs that are either zero or infinity.

The Mantel-Haenszel Method

The special case where no adjustment for confounding has been made for the odds ratios to be combined is of limited use in practice. It has the advantage of being easy to calculate when some of the a_i , b_i , c_i , or d_i are 0.

For study i , the number of exposed and unexposed cases and controls are given by:

	Case	Control
Exposed	a_i	b_i
Unexposed	c_i	d_i

The pooled estimate is calculated by

$$OR_{MH} = \frac{\sum_{i=1}^k a_i d_i / n_i}{\sum_{i=1}^k b_i c_i / n_i},$$

where n_i is the total number of subjects in the i^{th} study [11].

The estimate of variance most commonly used for deriving confidence intervals is

$$Var(OR_{MH}) = \frac{\sum_{i=1}^k P_i R_i}{2 \left(\sum_{i=1}^k R_i \right)^2} + \frac{\sum_{i=1}^k (P_i S_i + Q_i R_i)}{2 \left(\sum_{i=1}^k R_i \right) \left(\sum_{i=1}^k S_i \right)} + \frac{\sum_{i=1}^k Q_i S_i}{2 \left(\sum_{i=1}^k S_i \right)^2},$$

where

$$P_i = (a_i + d_i) / n_i,$$

$$Q_i = (b_i + c_i) / n_i,$$

$$R_i = a_i d_i / n_i, \text{ and}$$

$$S_i = b_i c_i / n_i.$$

Thus a $100(1-\alpha)\%$ confidence interval for $\ln(OR_{MH})$ is given by

$$\ln(OR_{MH}) \pm z_{1-\alpha/2} [Var(OR_{MH})]^{1/2} \quad [12].$$

Where covariates are adjusted for, in order to eliminate (or reduce) the bias caused by confounding, one method of achieving this is via stratification. Here an extended form of the Mantel-Haenszel odds ratio is used.

Suppose there are S strata, then the summary odds ratio is given by

$$OR_{MH(S)} = \frac{\sum_{j=1}^S a_j d_j / n_j}{\sum_{j=1}^S b_j c_j / n_j}$$

where a_j, \dots, d_j are equivalent to those above, but are at the stratum, rather than the study-level. Note that S can indicate an odds ratio from a study or statistically independent strata from within a study.

The standard error of the log of the summary odds ratio is given by

$$SE(\log(OR_{MH(S)})) = \frac{1}{\sqrt{2} \sum_{j=1}^S A_j} \left(\sum_{j=1}^S A_j B_j + OR_{MH(S)} \sum_{j=1}^S (B_j C_j + A_j D_j) + OR_{MH(S)}^2 \sum_{j=1}^S C_j D_j \right)^{0.5}$$

where

$$A_j = \frac{a_j d_j}{n_j},$$

$$B_j = \frac{1}{n_j} (a_j + d_j),$$

$$C_j = \frac{b_j c_j}{n_j}, \text{ and}$$

$$D_j = \frac{1}{n_j} (b_j + c_j).$$

A $100(1 - \alpha)\%$ confidence interval can then be constructed analogously as above. Note that $OR_{MH(S)}$ is a weighted average of the S stratum-specific odds ratios, given by

$$OR_j = \frac{a_j d_j}{b_j c_j}$$

with weights given by

$$w_j = \frac{b_j c_j}{n_j} \text{ [13].}$$

The Peto Method

This is essentially a modification of the Mantel-Haenszel method, and also has the advantage of being easy to calculate when some of the a_i , b_i , c_i or d_i above are 0.

However, it has been shown to produce serious underestimates when the odds ratio for the exposure effect is far from unity [14].

$$OR_{Peto} = \frac{\sum_{i=1}^k O_i - E_i}{\sum_{i=1}^k V_i},$$

where

$$V_i = E_i \frac{(n_i - n_{ei})(n_i - d_i)}{n_i(n_i - 1)},$$

with

n_i = number of subjects in study i ;

n_{ei} = number of exposed subjects in study i ;

d_i = number of deaths in the exposed and non-exposed groups, and

$$E_i = \frac{n_{ei} d_i}{n_i},$$

which is the expected number of deaths in the exposed group.

Thus a $100(1-\alpha)\%$ confidence interval is given by

$$\exp \left(\frac{\sum_{i=1}^k (O_i - E_i) \pm z_{1-\alpha/2} \sqrt{\sum_{i=1}^k V_i}}{\sum_{i=1}^k V_i} \right) \quad [15].$$

3.3.1.4 Combining correlation coefficients

The use of correlation coefficients in the occupational health field has been exclusively restricted to studies of psychosocial endpoints. Suppose there are k product moment correlation coefficients to be combined r_1, \dots, r_k . The most direct approach to estimating the overall correlation coefficient ρ is to derive a linear combination of the r_1, \dots, r_k . This is not recommended unless the sample sizes of all m data sets of size n_i are very large; the method usually used is to calculate the weighted average of z -transformed estimates yielding

$$p = \sum_{i=1}^k w_i z(r_i)$$

where

$$w_i = \frac{n_i - 3}{\sum_{j=1}^k (n_j - 3)}$$

and

$$z(r_i) = \frac{1}{2} \log \frac{1+r_i}{1-r_i}$$

(the equivalent transformation for ρ being

$$\zeta = \frac{1}{2} \log \frac{1+\rho}{1-\rho}).$$

A $100(1-\alpha)\%$ confidence interval for ρ can be obtained for ζ from

$$p \pm \frac{z_{1-\alpha/2}}{\sqrt{N-3k}},$$

where

$$N = \sum_{i=1}^k n_i$$

and thus for ρ by utilizing the transformation

$$\rho = y^{-1}(\zeta),$$

where

$$y^{-1}(x) = \frac{(e^{2x} - 1)}{(e^{2x} + 1)} \quad [16].$$

3.3.2 Random effects methods

This approach assumes that the studies are a random sample from a larger population of studies, and that there is a mean population effect size around which the underlying study-specific effect sizes vary. Expressed mathematically

$$Y_i = \mu + \delta_i + e_i,$$

where Y_i is the observed effect from study i , μ is the population effect size, δ_i is the deviation of the i^{th} study's underlying effect size from μ , and e_i is the sampling error.

Thus the δ_i are assumed to follow a normal distribution and represent the between study variance τ^2 [17], Thus

$$\text{var}(\delta_i) = \tau^2$$

and when

$$\text{var}(\delta_i) = 0,$$

the model reduces to the fixed effect model.

3.3.2.1 Combining generic effect measures (the DerSimonian and Laird method) by means of a weighted average approach

The mean weights from the k studies are defined as

$$\bar{w} = \frac{\sum_{i=1}^k w_i}{k}$$

where the w_i are as defined in the fixed effect approach and the variance of the weights to be

$$s_w^2 = \frac{1}{k-1} \left(\sum_{i=1}^k w_i^2 - k \bar{w}^2 \right).$$

Further, if

$$U = (k-1) \left(\bar{w} - \frac{s_w^2}{k \bar{w}} \right),$$

then the estimate of the variance due to inter-study variations in effect size is

$$\hat{\tau}^2 = \begin{cases} 0, & Q \leq k-1 \\ \frac{(Q - (k-1))}{U}, & Q > k-1 \end{cases}$$

where Q is the heterogeneity statistic (see Section 3.4).

Adjusted weights for the individual studies are given by

$$w_i^* = \frac{1}{\frac{1}{w_i} + \hat{\tau}^2},$$

and the overall estimate of effect size is

$$\bar{Y} = \frac{\sum_{i=1}^k w_i^* Y_i}{\sum_{i=1}^k w_i^*}.$$

A $100(1-\alpha)\%$ confidence interval can be estimated as

$$\bar{Y} \pm \frac{z_{1-\alpha/2}}{\sqrt{\sum_{i=1}^k w_i^*}} [18].$$

As for the fixed effect analysis, if Y_i is assumed to be a ratio (relative risk), then by putting

$$Y_i = \log RR_i$$

the mean relative risk on the log scale can be expressed as in 3.3.1.1 but with w_i replaced by w_i^* and the associated $100(1-\alpha)\%$ confidence interval can be expressed as before with

$$\text{var}(\overline{\log RR}) = \frac{1}{\sum_{i=1}^k w_i^*}.$$

Thus, assuming normality of RR on the log scale, a $100(1-\alpha)\%$ confidence interval can be expressed as before as

$$\overline{\log RR} \pm z_{1-\alpha/2} \sqrt{\text{var}(\overline{\log RR})} [8].$$

3.3.2.2 Combining Standardised Mortality Ratios (SMRs)

The random effects meta-SMR (meta-SMR_{RE}) has the same formula as that given in Section 3.3.1.2, but the weights w_i^* are given by

$$w_i^* = \frac{1}{\text{var}(\log SMR_i) + \hat{\tau}^2}$$

with

$$\hat{\tau}^2 = \max \left\{ 0, \frac{Q_{k-1} - (k-1)}{\sum_{i=1}^k w_i - \frac{\sum_{i=1}^k w_i^2}{\sum_{i=1}^k w_i}} \right\},$$

$$Q_{k-1} = \sum_{i=1}^k w_i \left[\log(SMR_i) - \overline{\log(SMR)} \right]^2,$$

and

$$\overline{\log(SMR)} = \frac{\sum_{i=1}^k w_i \log(SMR_i)}{\sum_{i=1}^k w_i}.$$

Thus an approximate $100(1-\alpha)\%$ confidence interval for the meta-SMR_{RE} is given by

$$\exp \left(\overline{\log SMR} \pm \frac{z_{1-\alpha/2}}{\sqrt{\sum_{i=1}^k w_i^*}} \right) [9].$$

3.3.2.3 Combining other effect estimators

Analogous risk estimates and associated $100(1 - \alpha)\%$ confidence intervals can be computed for the Mantel-Haenszel estimator (with and without adjustment for confounders), for the Peto estimator and for the combined correlation coefficients.

3.3.2.4 Maximum likelihood methods

Likelihood approaches exist under both the fixed effect and the random effects assumption. For the random effects case, with the assumption of the underlying effect parameter following a normal distribution, DerSimonian and Laird [18] proposed maximum and restricted maximum likelihood methods of parameter estimation (the latter being a modification to the former adjusting for the fact that the mean and the variance are estimated from the same data). Likelihood and restricted maximum likelihood equations that can be solved by iterative numerical methods have been presented [19].

For the specific case of combining odds ratios, maximum likelihood estimates exist under the fixed effect assumption, although they are difficult to compute, and are only useful when there are a large number of odds ratios to be combined. As well as unconditional estimates, estimates conditional on fixing the marginal totals in 2x2 tables are also available, and are superior to the Mantel-Haenszel estimate in terms of bias and precision [20].

3.3.2.5 Bayesian methods

Bayesian methods are becoming more widely applied in meta-analyses in occupational epidemiology. This is partly due to advances in computational methods. The Bayesian approach provides a framework in which the data and model parameters are considered random quantities and the likelihood function can be thought of as defining the

plausibility of the data given values of the model parameters. As the model parameters are unknown random quantities, prior distributions can be specified for them which may be based on actual evidence or on subjective *a priori* beliefs (which in practice can also be non-informative). The joint probability density function for all the parameters is then combined with the likelihood function to obtain the joint posterior probability density function.

Usually there are a large number of parameters in the model, but attention is usually restricted to a few of specific interest. In the context of a meta-analysis in occupational epidemiology, this is usually the effect size, μ , say. The marginal posterior density of μ is obtained by integrating the joint posterior density over all the other parameters; the more parameters the higher dimension the integral. Thus, applied use of Bayesian methods has involved computationally innovative procedures for this.

The method most popularly applied is the use of Markov Chain Monte Carlo methods for estimating the marginal posterior density and in particular the use of Gibbs sampling via the software WinBugs [21]. Caution has to be exercised to ensure that the Markov chain has properly converged and the sensitivity of the results to the choice of prior distribution and initial values is examined. Also problematic, as with any exercise in statistical modelling is determining the best model in the circumstances and whether it represents an adequate representation of the data.

The Bayesian framework allows for fixed effect and random effects methods to be pursued. Further extensions not practised much to date include the incorporation of informative prior distributions, perhaps based on the evidence of a pre-existing meta-analysis [22].

3.4 METHODS FOR ASSESSING THE PRESENCE OF BETWEEN-STUDY HETEROGENEITY

In meta-analyses the point estimates of effect size from the component studies are usually different. Of interest, is whether these differences can be attributed to random variation (the studies are estimating the same overall effect estimate) or whether they are estimating a distribution of effect estimates (and thus differences are not solely due to random variation). The strength of evidence against the underlying effects from the component studies being the same is conventionally assessed via a heterogeneity test, the null hypothesis being that there is a lack of heterogeneity. These tests are usually based on the Cochran χ^2 test and are acknowledged as not statistically very powerful [23].

3.4.1 Generic Cochran chi-squared test

The generic form of the test can be expressed as

$$Q_{k-1} = \sum_{i=1}^k w_i (Y_i - \bar{Y})^2,$$

where k , w_i , Y_i and \bar{Y} are defined in Section 3.3.1.1 and the null hypothesis of no heterogeneity would be rejected if Q exceeded the upper $100(1-\alpha)\%$ percentile of the χ^2 distribution with $k-1$ degrees of freedom [7]. Note, however, a lack of significance of this test does not imply homogeneity [24].

For multiple studies where the effect measure can be expressed in the form of a relative risk, a statistical test of the homogeneity assumption on the log scale is given by

$$Q_{k-1} = \sum_{i=1}^k w_i (\log RR_i - \overline{\log RR})^2,$$

where

$$\overline{\log RR} = \frac{\sum_{i=1}^k w_i \log RR_i}{\sum_{i=1}^k w_i} \quad [8].$$

3.4.2 SMRs

The statistic can be adapted for use with SMRs in an analogous way giving

$$Q_{k-1} = \sum_{i=1}^k w_i \left[\log(SMR_i) - \overline{SMR} \right]^2,$$

where

$$\overline{SMR} = \exp \left(\sum_{i=1}^k w_i \log(SMR_i) / \sum_{i=1}^k w_i \right) \quad [10].$$

3.4.3 Peto method

A heterogeneity test statistic based on the Peto method [15] is given by

$$Q_{k^*-1} = \sum_{i=1}^k \frac{(O_i - E_i)^2}{V_i} - \frac{\left[\sum_{i=1}^k (O_i - E_i) \right]^2}{\sum_{i=1}^k V_i}$$

with degrees of freedom equal to $k^* - 1$, where $k^* < k$ is the number studies with nonzero variances, and V_i defined as in Section 3.3.1.3.

3.4.4 Correlation coefficients

Where the statistics being combined are correlation coefficients, a heterogeneity test statistic is given by

$$Q_{k-1} = \sum_{i=1}^k w_i [z(r_i) - \bar{z}]^2 ,$$

where $z(r_i)$ and w_i are as given in Section 3.3.1.4, and

$$\bar{z} = \frac{\sum_{i=1}^k w_i z(r_i)}{\sum_{i=1}^k w_i} .$$

3.5 METHODS FOR EXPLORING BETWEEN-STUDY HETEROGENEITY

It is now generally agreed that meta-analysis can and should go further than simply producing overall summaries of effect [19]. Thus the main focus of meta-analyses should be an exploration of the reason why the different component studies within a meta-analysis produce different effect estimates (or at least more than can be explained simply by sampling variation).

3.5.1 Graphical methods

A graphical examination of heterogeneity is normally carried out in addition to a test based on the Cochran χ^2 test. A forest plot, which consists of the individual study effect sizes together with their associated confidence interval being plotted on the same graph, provides a useful informal comparison of the effect sizes from the different component studies, usually at the exploratory stage of a meta-analysis. A useful feature is that the plotting symbol can be made proportional in size to the inverse of the variance of the effect estimate, to provide a visual representation of the relative weight of each study.

3.5.1.1 *Plots of normalised scores*

Here the standardised residuals for each study can be calculated as

$$z_i = \frac{Y_i - \bar{Y}}{\sqrt{\text{var}(Y_i)}},$$

where Y_i and \bar{Y} are as in Section 3.3.1.1, when under the null hypothesis of only random differences between the studies, a histogram of the z-scores would have an approximately normal distribution with mean 0 and variance 1. It is therefore usual to superimpose on the histogram a normal distribution with mean 0 and variance 1. Large absolute z-scores can signal important deviations from the average result, thus providing an indication of which studies are responsible for large contributions to the heterogeneity test statistic [15].

3.5.1.2 *Radial plots*

These plots consist of the outcome from each study divided by the square root of its

variance $\left(\frac{Y_i}{\sqrt{\text{var}(Y_i)}} \right)$ plotted against the reciprocal of the standard error $\left(\frac{1}{\sqrt{\text{var}(Y_i)}} \right)$.

The position of each study in relation to the horizontal axis provides an indication of the weight associated with each study (the smaller the weight the larger the standard error and the further it will be from the abscissa). Deviations from the simple linear regression line constrained to fit through the origin indicate possible outliers contributing considerably to the between study heterogeneity [25]. Subgroups of studies may be highlighted by the use of colour.

3.5.2 Subgroup analysis – stratification by study characteristics

The main type of subgroup analysis possible in meta-analyses in occupational epidemiology is that which investigates subsets of workers defined by study

characteristics (e.g. country, industry, occupational group). It is important that the subgroups are clearly defined (preferably defined *a priori*). Regression models are required, however, if there are many different characteristics that need to be examined at the same time.

3.5.3 Fixed effect and random effects regression models

There are two types of regression model, the fixed effect or meta-regression model and the random effects regression or mixed model (so called because it includes fixed and random terms).

The meta-regression model is used when all the variation, other than sampling error, is explainable by the covariates included in the model (usually study characteristics). A mixed model is used when the covariates only partly explain the variation or heterogeneity and a random effects term is employed to explain the remainder. Clearly in practice, not all the variation (other than sampling error) will be explained by the covariates and so a choice between the regression approaches will need to be made.

A mixed model can be expressed as

$$Y_i = \beta_0 + \sum_{j=1}^P \beta_j x_{ij} + \delta_i + e_i ,$$

i.e.

$$Y_i = \mu + \delta_i + e_i \text{ (see section 3.3.2)}$$

where:

$$Y_i = \log(RR_i)$$

and

$$\mu = \beta_0 + \sum_{j=1}^P \beta_j x_{jp};$$

And the e_i are the sampling errors as in Section 3.3.1; the δ_i are the deviations of the i^{th} study's underlying effect size from μ as in Section 3.3.2; the x_{ij} are the values of the predictor variables for the i^{th} study; and the β s are the $p+1$ unknown regression parameters to be estimated, each of which estimates the influence of a study characteristic on the magnitude of the outcome of interest. As for the weighted average approach, if

$$\text{var}(\delta_i) = 0,$$

then the model reduces to the fixed effect meta-regression model.

The usual standard errors for the parameter estimates from regression models can be computed and thus associated $100(1-\alpha)\%$ confidence intervals can also be derived.

The explanatory variables are characteristics of the studies and the magnitude of the β s indicates the potential impact on the overall meta-RR. The testing of the usual assumptions underlying the regression models, as well as the testing of model adequacy and for influential risk estimates, need to be carried out before drawing conclusions.

3.6 PUBLICATION BIAS

Identification of all possible epidemiological investigations is important, since identification of only a subset has the potential to lead to biased conclusions. Complete ascertainment of all possible studies is unlikely to happen in practice, since the practice of writing up or submitting for publication is related to whether or not the study findings are positive. In such instances the ensuing bias in the overall effect estimate caused by this is known as publication bias.

3.6.1 Funnel plots (identification of publication bias)

Detection of publication bias can be carried out informally using a visual method based on a funnel plot. A funnel plot is simply a scatter plot of the effect estimates of interest from individual studies against some measure of each study's size or precision. An asymmetric graph, usually caused by a lack of small studies with low relative risk or an exaggeration of effect sizes in small studies of low quality, suggests the presence of publication bias. Funnel plots are at their most useful in the presence of studies with a range of sizes [26], as a spread of observations in the vertical direction will make the absence, for example, of small null studies easier to spot.

Other, more formal, methods exist for identifying publication bias. For example, a method exists where the association between effect estimates and their variances is examined via a rank correlation test [27]. The normalised test statistic is compared to the standardised normal distribution. This Begg and Mazumdar test is usually carried out alongside the funnel plot, and is not considered powerful when dealing with a small number of studies. Other methods include a linear regression test [26] in which the effect estimate divided by its standard error, producing a standard normal deviate, is regressed against the estimate's precision (inverse of standard error), and a 'trim and fill' method [28] also exist.

3.6.2 The file-drawer method (measurement of extent of publication bias)

This is a method that attempts to answer the question: "how many studies have been conducted averaging a null result are needed to bring the overall effect estimate to not statistically significant?"

The normal z-scores (Z_i) associated with the p-values observed for each study can be combined to produce an overall z-score:

$$Z = \frac{\sum_{i=1}^k Z_i}{\sqrt{k}},$$

where k is the number of effect estimates included in the meta-analysis. If k_0 is defined as the number of unpublished studies with an average observed effect size of 0 that would be needed to make the overall z-score Z not statistically significant, then it can be estimated as:

$$k_0 > -k + \frac{Z^2 k}{Z_{1-\alpha/2}^2}.$$

Once calculated a judgment has to be made about whether k_0 represents a realistic number of studies that have been “filed away”. Further, it has been suggested that $N = 5k + 10$ should be considered a fail-safe such that if k_0 is less than this level then one must have doubts about the validity of the meta-analysis [29].

3.6.3 Brief overview of other methods for dealing with publication bias

If feasible, then efforts should be made to track down studies missing from a meta-analysis, although this will usually mean the incorporation of results that have not been through the peer-review process. In practice, this is usually too difficult. One of the most straightforward methods to adjust for publication bias in a meta-analysis is to include only the largest studies in a meta-analysis. This involves an arbitrary cut-off for deciding what constitutes a large study. Another possibility is to carry out a cumulative meta-analysis in which studies are combined sequentially in order of precision.

Approaches other than the file drawer method exist for estimating the number of unpublished studies [30]. The ‘trim and fill’ method mentioned in Section 3.6.1 provides a method for adjusting for publication bias. In addition a method has been

developed based on a random effects meta-analysis model via a sensitivity analysis approach [31].

3.7 OTHER META-ANALYTICAL METHODS

This section includes details of other methodological methods not already covered that may be important for the occupational setting.

3.7.1 Methods for combining dose-response data

This is an area of meta-analysis that is very important in the occupational setting, particularly in the regulatory risk environment where decisions have to be made about whether or not an association represents a causal one and what constitutes a “safe” occupational exposure limit. The following contains only a very brief introduction to the topic.

Combining dose response data involves the pooling of correlated estimates to compute regression slopes across different occupational exposure categories. Methods have been proposed that account for these correlations but only require summary estimates of the marginal data from the component studies. Some of the methods can also be extended to allow for non-linear increases in risk with exposure [32].

3.7.2 Methods for assessing and adjusting for study quality

If a general formula for the fixed effect weighted average effect size over the k studies is

$$\overline{RR} = \frac{\sum_{i=1}^k w_i RR_i}{\sum_{i=1}^k w_i},$$

then a quality adjusted weight may be defined as

$$w_i^{\#} = w_i q_i$$

where q_i is the quality score for study i and

$$w_i = \frac{1}{\text{var}(RR_i)}.$$

Thus the score for each study may be incorporated into the overall average effect size by

$$\overline{RR} = \frac{\sum_{i=1}^k w_i^{\#} RR_i}{\sum_{i=1}^k w_i^{\#}},$$

whence a $100(1-\alpha)\%$ confidence interval is given by

$$\overline{RR} \pm z_{1-\alpha/2} \sqrt{\frac{\sum_{i=1}^k w_i^{\#} q_i}{\left(\sum_{i=1}^k w_i^{\#}\right)^2}} \quad [33].$$

This approach can be extended to a random effects approach by using the DerSimonian and Laird approach. Note, however, that the use of quality scores in meta-analysis remains controversial [34] and these are discussed in section 3.8.

3.7.3 Sensitivity analysis

Sensitivity analyses can play an important part in determining the sensitivity of the findings of the results of a meta-analysis to the approach chosen. This will be particularly important in the context of statistical modelling where for example assumptions of normality (or, more usually, normality on the log scale) for effect

measures and independence of studies cannot be easily tested. Other important issues, such as the effect of the choice of inclusion/exclusion criteria or of the choice of quality scores may also be important to investigate. It is prudent therefore to try to demonstrate consistency of conclusions from a variety of approaches. A lack of consistency would provide a focus for areas for further investigation.

3.8 SOME EXAMPLES OF THE CONTROVERSIAL ASPECTS OF THE METHOD

Some aspects of the meta-analysis method are of central importance or specifically relevant in occupational epidemiology. For example, the summary Standardised Mortality Ratio (SMR) or meta-SMR has been used because of its simplicity, ease of interpretation, its preservation of the adjustments in the original studies, and its resemblance to the analyses in the original studies [35]. However, it has also been suggested that the validity of the meta-SMR is uncertain, even though it has the virtue of weighting the various SMRs (from the different studies) by the number of expected deaths and is (therefore) of some value [36]. Furthermore, it has been forcefully stated that SMRs from cohorts with different age structures cannot be validly compared (although there may be circumstances when this is not so) and that this problem is generally ignored in meta-analysis in occupational epidemiology on the assumption that the age distributions of the cohorts to be combined are approximately equal [9]. (This issue also applies to subgroup analyses within cohort studies.) Another major difficulty faced by the meta-analyst in the occupational health field is the differences in exposure measurements between studies [37]. For example, in cohort studies where exposures are classified by a broad definition such as a job title, not all of those workers will have the same exposures [38] and combining those studies with different exposures would tend to dilute the effects of the relevant exposures and bias the risk estimates towards the null [39, 40].

3.9 RECENT METHODOLOGICAL DEVELOPMENTS (SINCE THE ORIGINAL REVIEW)

This section briefly summarises some recent developments in meta-analysis methodology that are relevant for practice in occupational epidemiology.

In their 2008 paper, Sutton and Higgins recognized that the need for research and practice to be based on the totality of relevant and sound evidence had been increasingly recognized and that the impact of meta-analysis in this had grown enormously [41]. They outlined how emphasis has been placed on (i) heterogeneity and random-effects analyses; (ii) special consideration in different areas of application; (iii) assessing bias within and across studies; and (iv) extension of ideas to complex evidence synthesis. They conclude that any new method should be backed up by appropriate simulation work, and that non-statistical considerations can have an important role in the choice of meta-analysis model.

Reitsma *et al* pointed out that several advances have been made in the methods used in performing meta-analyses of diagnostic test accuracy studies, most notably how to assess the methodological quality of diagnostic test accuracy studies and the paired measures of test accuracy (bivariate meta-regression model of sensitivity and specificity) [42].

The multivariate random effects model is a generalization of the standard univariate model. Multivariate meta-analysis is becoming more commonly used and the techniques and related computer software, although continually under development, are now in place. Jackson *et al* have described the areas of application that multivariate meta-analysis has found, the methods available, the difficulties typically encountered and the arguments for and against the multivariate methods. They concluded that the multivariate methods can be useful, and in particular can provide estimates with better

statistical properties, but also that these benefits come at the price of making more assumptions which do not result in better inferences in every case. Thus although there is evidence that multivariate meta-analysis has considerable potential, it must be even more carefully applied than its univariate counterpart in practice [43].

Higgins in a Lancet editorial explained network meta-analysis as applied to clinical trials. In the absence of a direct head-to-head comparison of two treatments, results of separate studies of those treatments need to be drawn on. He explained that a naïve indirect comparison could be made by comparison of outcomes of people receiving one treatment in one study with those receiving the other treatment in a different study, but that this approach could easily give the wrong answer, because it doesn't account for key differences between studies and loses the advantage gained by the use of randomization. Network analysis combines results from all studies simultaneously, drawing on both direct comparisons within studies and indirect comparisons across studies via common reference treatments. In a standard meta-analysis, studies need to be sufficiently similar to each other for the combined result to be meaningful. In a network meta-analysis – see for example Siontis *et al* [44] – studies need to be sufficiently similar in ways other than the particular choice of treatments being compared. The method allows estimation of both heterogeneity in the effect of any given treatment and inconsistency in evidence from different pairs of treatments [45]. The method is closely related to multivariate meta-analysis and it is possible to see how this approach could be adapted for use in occupational epidemiology.

Wei and Higgins published an extension of the method of Bayesian multivariate meta-analysis to cover the situation when there are more than two outcomes of interest. Their approach includes marginal modelling of data, to account for the fact that not all relevant studies necessarily publish data on all the outcomes of interest [46]. There are

currently very few examples if any of multivariate meta-analyses, whether Bayesian or frequentist in the occupational epidemiology literature.

An assessment of bias in randomized controlled trials indicated that there was heterogeneity between trials in terms of the baseline characteristics of the study populations to be combined and this resulted in bias being introduced into the meta-analysis [47]. This is clearly an issue that is extremely important within the context of meta-analyses in occupational epidemiology and one which has yet to be fully addressed in examples in the literature.

An interesting recent development has been the introduction of an inverse variance quasi-likelihood based alternative to the random effects model, the so called IVhet model [48]. The paper suggests that the IVhet model is an improved alternative to the random effects model for meta-analysis of heterogeneous studies. It shows that the known issue of the underestimation of the statistical error and thus spuriously overconfident estimates with the random effects model can be resolved by use of an estimator under the fixed effect model assumption with a quasi-likelihood based variance structure. The authors claim that extensive simulations have confirmed that the estimator retains a correct coverage probability and a lower observed variance than the random effects model estimator, regardless of the level of heterogeneity. They recommended that the IVhet model be used in practice instead of the fixed effect and random effects models. This is an interesting very recent development. There are no examples yet where this has been utilized in the occupational epidemiology setting.

A systematic review of simulation studies comparing the performance of different estimation methods for the estimate of between-study heterogeneity in random effects meta-analysis has recently stated that the DerSimonian and Laird method is negatively biased when heterogeneity is moderate to high and for most studies recommended an

alternative [49]. Three of the studies they reviewed recommended the Paule-Mandel method [50] as being simple to implement and being less biased than the DerSimonian and Laird method, performing well with dichotomous and continuous outcomes. The authors of the review provisionally recommended this method, but cautioned that further simulation studies are required before firmer conclusions could be drawn [49].

The Cochrane collaboration has produced a tool to assess bias in randomized controlled trials (see

[http://handbook.cochrane.org/chapter_8/8_assessing_risk_of_bias_in_included_studies.](http://handbook.cochrane.org/chapter_8/8_assessing_risk_of_bias_in_included_studies.htm)

[htm](http://handbook.cochrane.org/chapter_8/8_assessing_risk_of_bias_in_included_studies.htm)). However, a recent review of its use in practice has mainly been in a non-recommended way and it has been recommended that its structure be improved and more focused guidance be produced[51].

A tool also now exists for assessing the quality of meta-analyses [46].

A recent paper has indicated that the inverse variance methods perform poorly when the data contain zeroes in either control or intervention arms in clinical trials and recommended that methods based on Poisson regression with random effects terms for the variance components are very flexible and offer substantial improvement [52].

4 META-ANALYSIS FOR RARE DISEASE OUTCOMES: WHAT TO DO WHEN THERE ARE NO EXPOSED CASES?

4.1 INTRODUCTION

The issue of sparse and missing data is an issue of major concern in randomized controlled clinical trials. Meta-analyses including the calculation of confidence intervals in statistical software e.g. Stata [53], are often carried out on the log scale as the distributions are more likely to be normal and their results transformed back to the original scale. However those studies that have a zero relative risk are excluded from such analyses because $\log 0$ is undefined. It is clear, however, that such studies are potentially important because they contain important statistical information, especially when the disease under consideration is rare, and so should be included in the analyses.

This chapter aims to explore possible approaches available to the meta-analyst when a non-trivial number of studies of a rare disease outcome have an observed relative risk equal to zero, and to make recommendations about the optimum approach in such circumstances. Studies of formaldehyde and nasopharyngeal cancer and sinonasal cancer will be used as a basis for this exploration.

4.2 SUMMARY OF THE LITERATURE ON THE META- ANALYSIS OF SPARSE DATA SETS

A search of the Science Direct and PubMed databases was carried out using the search terms ‘meta’ and ‘analysis’ combined with ‘sparse’ or ‘zero’ in order to find relevant papers that deal with sparse data sets. By 2007, there were fewer than 10 papers have been published on this issue. Previously authors have concentrated on issues relating to

the pooling of data from controlled clinical trials, and thus the pooling of odds ratios.

These papers are summarised below.

Sankey *et al*, 1996

Sankey *et al* explored two possible approaches for carrying out meta-analyses of sparse data sets that could be expressed in the form of a 2x2 contingency table using simulations: a corrected method in which one half was added to each cell, and an uncorrected method. These methods were compared over a range of sparse data situations in terms of coverage rates using three summary statistics: the Mantel-Haenszel odds ratio and the DerSimonian and Laird approach using the odds ratio and the rate difference. The uncorrected method performed better only when using the Mantel-Haenszel odds ratio with very little heterogeneity present. For all other sparse data applications, the method with the continuity correction performed better and is the one recommended by authors for use in meta-analyses of similar scope [54].

Austin *et al*, 1997

Austin *et al* contended that although there are a variety of procedures available for combining effect measures across epidemiologic studies, none of the methods provided an overall effect estimate when the data were sparse within studies and come from different study designs. In their paper they discussed the statistical relationship between case-control studies and two types of follow-up studies. Their method relies on the data from follow-up studies being able to be expressed in the form of a 2x2 contingency table. They developed an exact methodology for combining results across study designs. They also derived Mantel-Haenszel-type formulae for summarising results across studies, and illustrated the techniques with data from studies of breast implants and connective tissue disease. The authors identified the limitations of their approach in that: continuous data needed to be categorised; prudence was needed in the combining

of results from different study designs; discrepancies in exposures and other study characteristics affect comparability; and estimates adjusted for confounders by stratification or regression could not be used, but that in practice adjusted and unadjusted estimates of effect size were often quite close [55].

Sutton *et al*, 2002

Sutton *et al* examined this problem in the context of binary data that could be expressed in the form of a 2x2 table in the context of rare adverse event data in clinical trials. The risk measures considered were:

- The risk difference
- The risk ratio of beneficial and harmful outcomes
- The odds ratio
- The numbers needed to harm (does not have properties required for meta-analysis, but can be derived from other summary statistics)

They stated that the choice between the different effect measures was far from straightforward. The risk ratio and the odds ratio are approximately equal for rare events and so the choice comes down to an absolute or relative comparator. They eventually concluded that they preferred the risk difference due to its direct interpretability and potential inclusion of studies in which zero events occurred [56].

Sweeting *et al*, 2004

Sweeting *et al* compared via simulation the performance of different classical and Bayesian meta-analysis methods for pooling odds ratios when applied to sparse event data with emphasis on the use of continuity corrections [57]. The continuity corrections used were either a constant or two alternatives: one based on a function of the reciprocal of the opposite group arm size; and the other on an empirical estimate of the pooled

effect size from the remaining studies in the meta-analysis. A variety of scenarios were simulated whilst varying the ratio of the study arm sizes. The methods employed were:

- Inverse-variance-weighted fixed effect model [58]
- Mantel-Haenszel summary odds ratio [11]
- Peto's method for combining odds ratios [15]
- Logistic regression using iterative weighted least squares
- Logistic regression via a Bayesian model [22]

Sweeting *et al* found that the Mantel-Haenszel summary estimates using the alternative continuity correction factors gave the least biased results for all group size imbalances. Logistic regression was unbiased for all scenarios and gave good coverage probabilities. The Peto method provided unbiased results for balanced treatment groups, but the bias increased with increasing ratio of study arm sizes. The Bayesian fixed effect model provided good coverage for all group size imbalances. The two alternative continuity corrections outperformed the constant correction factor in nearly all situations. The inverse variance method always performed badly, irrespective of the continuity correction used. They concluded that a sensitivity analysis using several methods and continuity correction factors is required for routine practice [57].

Bradburn *et al*, 2007

Bradburn *et al* evaluated, via simulation, the performance of 12 methods, readily available in software, for pooling rare events and considered estimability, bias, coverage and statistical power [59]. Their simulations were based on data sets from three case studies with between five and 19 trials using baseline event rates between 0.1 and 10 per cent and risk ratios of 1, 0.75, 0.5 and 0.2. The methods they employed were:

- Odds ratios (continuity corrections of adding 0.5 to each 2x2 cell were made only where absolutely necessary and feasible):
- Inverse variance - fixed effect (variance of individual trial ORs calculated using Woolf's method [60])
- DerSimonian and Laird - random effects [18]
- Mantel-Haenszel with and without zero-cell correction (their pooled variance estimators being the unconditional product binomial estimators [12, 13, 61])
- Mantel-Haenszel with continuity correction (their pooled variance estimators being the unconditional product binomial estimators [12, 13, 61])
- The Peto 'one-step' method [15]
- Exact stratified method [62]
- Maximum likelihood from logistic regression
- Risk differences
- Inverse variance - fixed effect (variance of risk difference estimated using a Normal approximation [63])
- DerSimonian and Laird - random effects [18]
- Mantel-Haenszel (their pooled variance estimators being the unconditional product binomial estimators [12, 13, 61])

In addition two unstratified marginal methods were evaluated [64, 65].

Bradburn *et al* found that most of the commonly used meta-analytical methods were biased when data were sparse and that the bias was greatest in inverse variance and DerSimonian and Laird odds ratio and risk difference methods, and for the Mantel-Haenszel odds ratio with a 0.5 continuity correction. Risk difference meta-analytical

methods tended to show conservative confidence interval coverage and low statistical power at low event rates. At event rates below one per cent, the Peto one-step method was the least biased and most powerful method, and provided the best confidence interval coverage, provided there was no substantial imbalance between treatment and control group sizes within trials, and treatment effects were not exceptionally large. In other circumstances the Mantel-Haenszel odds ratio approach in the absence of use of continuity corrections, logistic regression and the exact method performed similarly to each other, and were less biased than the Peto method [59].

In summary, for studies that can be expressed in the form of a 2x2 table, there is no clear consensus on how the meta-analysis should be carried out. There appears to be a difference of opinion on whether meta-analyses of odds ratios are to be preferred to meta-analyses of risk differences. In meta-analyses of odds ratios, caution is advocated when continuous data has been categorised, a continuity correction should be used and a sensitivity analysis should be carried out using several approaches to examine the robustness of any conclusions. None of the methods previously applied include the appropriate use of adjusted relative risks; it is just argued that unadjusted relative risks provide a good approximation to adjusted relative risks. None of the literature deals specifically with methods that could be used in meta-analyses of studies used in occupational epidemiology.

4.2.1 More recent methodological developments

Friedrich *et al* demonstrated that inclusion of trials arms with no events in a meta-analysis of randomized controlled trials could have a crucial impact on any inferences that may be drawn [66]. They used Mantel-Haenszel inverse variance methods, exact statistical methods, Bayesian analysis with non-informative priors and analyses of

additionally related outcomes to conduct alternative analyses to those that had previously been published.

A random effects meta-analysis approach has been proposed in which the approximately normal within-study likelihood is replaced by the appropriate exact likelihood [67], leading to a generalized mixed model that can be fitted using standard statistical software. It can also be fitted in situations where studies have zero exposed cases.

Cai *et al* proposed a random effects model based on the Poisson distribution [68].

Bohning *et al* reminded meta-analysis practitioners that meta-analysis of rare event studies had recently become a topic of controversy and debate [69]. They argued and demonstrated that the occurrence of zero events in clinical trials or cohort studies, even if zeroes occur in both arms is less problematic, at least from a statistical perspective, if the available statistical tools are applied in an appropriate way. In particular that cautioned against the exclusion of studies with zero events from meta-analyses. They focused on Mantel-Haenszel techniques, mixed Poisson regression and related regression models.

Bayesian estimators of the treatment effect and the heterogeneity parameter, as well as hypothesis testing methods based on Bayesian model selection procedures [70], have been compared by simulation to moment-based approaches under a random effects model [71]. This last method deals with zero events via a simple average estimator although the bias will tend to increase as the event rate gets lower and the number of studies with zero events gets higher.

4.3 SYSTEMATIC PRESENTATION OF OPTIONS

The aim of this section is to focus on options of relevance in occupational epidemiology, and of cohort studies in particular. It is rare that there will be no exposed cases of a rare disease in a type of study other than a cohort study. Clearly there will be studies other than occupational cohort studies that will contain data of relevance and so a couple of alternative methods are presented to allow their inclusion. In rare circumstances where there a large number of informative studies that can be included in the meta-analysis, an exploration of sources of heterogeneity via meta-regression might be undertaken.

There are several approaches for inclusion of studies with zero relative risk in the estimation of an overall mean, which can be applied to cohort or PMR-type studies:

- Carry out analyses on the log scale
 - ignoring studies with $O = 0$
 - by making adjustments to the data so that they will not introduce any important bias into the analysis, but in a way that allows the studies with $RR = 0$ to be included in an analysis on the log scale; this may involve adding a constant to the O and E for the studies with $O = 0$
- Carry out the analysis on the original (unlogged) scale; and
- Carry out analyses using transformations other than logs.

If the analysis includes studies other than cohort or PMR-type studies, then such studies can be included by:

- Deriving a pseudo expected number of cases and applying the method for cohort studies

- Analysing them separately on the log scale, with the results presented separately from those of the cohort studies or (if appropriate) for the end results of the two analyses to be combined to produce a single overall estimate of effect.

If there are sufficient studies to make exploration of heterogeneity meaningful, carry out a meta-regression on the original scale (Poisson for fixed effect; negative binomial for random effects). In the form of null models (no fitted covariates), these could be used as additional options to obtaining single pooled estimates of meta-relative risk.

These options are expanded upon in the following sections.

4.3.1 Analyses on the log scale

As mentioned in section 4.1, analyses of relative risks in meta-analyses are often carried out on the log scale. These are backed up by an observation that much of the available statistical software has been written for analyses on the log scale, for example, the ‘metan’ routine for Stata [53]. The motivation for using the log scale is that it helps to symmetrise the distributions of random variables and thus make them more normal [9]. Because the log of the relative risk is essentially a sum of independent random variables, it should converge to normality under the central limit theorem, allowing an approximate 95% confidence interval to be calculated [9]. However, there remains a problem when a non-trivial number of cohort studies involving a relatively rare event such as nasopharyngeal cancer or sinonasal cancer are studied, since there is a high probability that more than a few studies will contain no observed cases of interest and thus the relative risk equals zero and so the log of the relative risk is undefined. In general, approaches to meta-analysing RRs involve taking logarithms and so at first sight are not possible here. There remain two obvious options: ignore the studies with no observed cases; and finding a way to include them by adding small constants to

avoid having to take logs of zero. The additional option of Poisson or negative binomial regression is also possible, and has been explored.

4.3.1.1 Ignoring studies with zero observed cases

This approach is often adopted subconsciously by meta-analysts who simply use the available software that performs the analyses on the log scale by excluding those studies with zero observed cases. The analysis proceeds using the methods outlined in Chapter 3. In practice, the upward bias in meta-RR is likely to be small, unless there are an informative number of studies with a zero relative risk. This can occur in occupational cohort studies of rare cancers.

4.3.1.2 Adding a constant to avoid taking logs of 0

The attempts to address this problem in the literature have almost exclusively concentrated on situations in which the data are from a controlled clinical trial and when the data from the component studies can be expressed in the form of a 2x2 contingency table in order to derive odds ratios. Thus approaches to address the problem of zero observed (exposed) cases have centred on the use of continuity corrections to prevent having to take logs of zero. The use of a continuity correction in such circumstances is based on the approximation of a discrete distribution by a continuous one. It has been argued that the use of a continuity correction of 0.5 makes the estimate of the mean unbiased [72]. The use of smaller values as low as 10^{-8} have also been proposed for contingency tables in general [73], but none has been used in the meta-analysis context [57].

Thus, as before, if Y_i is assumed to be a ratio (a relative risk), then putting

$$Y_i = \log \frac{O_i}{E_i}$$

can be modified to $Y_i = \log \frac{O_i + k}{E_i + k}$, so

$$Y_i = \log RR_i^*,$$

where

$$RR_i^* = \frac{O_i + k}{E_i + k}.$$

Thus a small constant k can be added to all the observed and all the expected values, to both the observed and expected values when $O = 0$, to the observed values only, or to the observed only when $O = 0$. In practice, it seems intuitive, that k should be added to both O and E for each study and that k should be small to minimise the resultant bias in the $\log RR$.

The mean relative risk on a log scale can then be expressed in the usual way as in Section 3.3.1.1 as

$$\overline{\log RR^*} = \frac{\sum_{i=1}^k w_i \log RR_i^*}{\sum_{i=1}^k w_i},$$

where

$$w_i = \frac{1}{\text{var}(\log RR_i^*)}.$$

Clearly, this approach assumes that the RR can be expressed as a ratio of observed to expected number of cases.

4.3.2 Analyses on the original scale

There exist approaches that involve analyses on the original (unlogged) scale, provided the relative risks can be expressed in the form O/E. They assume that the variance of the individual study effect sizes can be estimated. The fixed effect weighted average approaches can be applied and appropriate confidence intervals can be calculated using exact methods. Although a central estimate can be calculated from the random effects method, the calculation of appropriate confidence intervals requires assumptions of normality for the estimates of between and within study variance.

If the number of observed and expected events are available from the individual cohort or PMR-type studies, then as outlined in Chapter 3, the meta-SMR can be calculated simply by using

$$meta - SMR = \frac{\sum_{i=1}^k O_i}{\sum_{i=1}^k E_i} \quad (\text{Herbold, 1993}),$$

(referred to previously as meta-SMR₂) which is algebraically equivalent to

$$meta - SMR = \frac{\sum_{i=1}^k w_i SMR_i}{\sum_{i=1}^k w_i},$$

where

$$w_i = E_i.$$

Exact confidence intervals can be calculated via the Poisson distribution, for example by using its relationship with the chi-squared distribution [74].

Now,

$$\text{var}(SMR_i) = \text{var}\left(\frac{O_i}{E_i}\right) = \frac{1}{E_i^2} \text{var}(O_i) = \frac{E_i}{E_i^2} = \frac{1}{E_i}$$

A simple extension to incorporate random effects can be derived. Here

$$meta - SMR = \frac{\sum_{i=1}^k w_i^* SMR_i}{\sum_{i=1}^k w_i^*},$$

where

$$w_i^* = \frac{1}{\frac{1}{w_i} + \tau^2},$$

$$\tau^2 = \max \left\{ 0, \frac{Q - (k-1)}{\sum_{i=1}^k E_i - \left(\frac{\sum_{i=1}^k E_i^2}{\sum_{i=1}^k E_i} \right)} \right\},$$

$$Q = \sum_{i=1}^k E_i (SMR_i - \overline{SMR})^2,$$

and

$$\overline{SMR} = \frac{\sum_{i=1}^k O_i}{\sum_{i=1}^k E_i} = meta - SMR.$$

This approach is outlined in a Stata technical bulletin [75].

For fixed effect analysis, the meta-SMR is equivalent to Poisson regression with no covariates. For random effects analysis, the meta-SMR would be expected to be similar to random effects Poisson regression which, when the errors are distributed according to the gamma distribution, is equivalent to negative binomial regression. This approach can be extended to studies other than cohort studies, such as case-control studies, by regarding the expected number of cases as though a “pseudo” expected number of cases from a cohort study (see section 4.3.4). Alternatively, for non-cohort studies, better estimates for standard errors on the original scale (e.g. to take account of extra uncertainty from controls in case-control studies) may be generated from the standard error on the log scale using the delta method.

Thus:

$$SE(RR) = SE(\log RR) * RR.$$

Therefore for analyzing studies that are a mixture of cohort and non-cohort studies, rather than the meta-SMR, a standard meta-analysis method involving RRs and their SEs is required. Meta-analyses on the original scale can, in theory, be extended as a form of meta-regression by including study characteristics as explanatory variables in the model.

4.3.3 Transformations other than taking logarithms

A variety of transformations are available for normalising a data set, the one most usually used in meta-analysis being that of taking logarithms. In addition, a square root transformation can be used. In one meta-analysis of formaldehyde and respiratory cancer [76], the authors made use of a square root transformation:

“However a number of studies had no cases of nasopharyngeal or nasal cancers, which precluded the logarithmic transformation typically utilized in meta-analyses formulae. In these cases, we used a square-root transformation of the lower and upper confidence intervals to derive weights for the individual studies and calculated the mRR by summing the products of the study weights and the RRs from the individual studies.”

It is not clear precisely how this transformation was implemented. It can be speculated that the square root transformation may have been applied to all studies when there was at least one study with a zero SMR or only to the studies with a zero SMR. Thus, the square root transformation implemented in this chapter may not be an implementation of the method used by Collins *et al.*

For cohort studies, the variance of the square root of a Poisson random variable can be approximated by $1/4$, the approximation improving as the mean increases [77]. So

$$\text{var}(\sqrt{RR}) = \text{var}\left(\frac{\sqrt{O}}{\sqrt{E}}\right) = \frac{1}{E} \text{var} \sqrt{O} \approx 0.25/E .$$

4.3.4 Dealing with studies other than cohort studies

Odds ratios from case-control studies can be incorporated into analyses on the original scale (4.3.2), provided an associated standard error for the odds ratio can be derived.

The standard error may be estimated directly on the original scale or from the estimated standard error on the log or square root scale via the delta method. The delta method has the advantage that it does not require the data to be in O/E format in order for the study to be used in a meta-analysis. Deriving the standard error directly on the original scale requires the data to be in O/E format; this can be achieved using a pseudo expected number of cases (pseudo $E = O/OR$), with the additional assumption that

pseudo expected values from studies with adjustment for confounding factors may be estimated by a pseudo expected value from an unadjusted estimate [53, 78]. Empirical evidence suggests that the difference between unadjusted odds ratios and confounder-adjusted odds ratios in practice is not large. The delta method [78] can be used to transform estimates from the log scale to the original scale; it is implemented in the Stata software [53] and so is judged a reliable approximation.

Thus, irrespective of whether the analyses are carried out on the original, square-root or log scales, there are two main approaches that can be adopted. The first is to combine the results from non-cohort studies with those of cohort studies to derive an overall estimated mean relative risk. The second is to carry out meta-analyses separately for each study type and if appropriate to combine these risk estimates together to obtain an overall estimate of relative risk.

4.3.5 Meta-regression

Standard meta-regression extends fixed effect and random effects meta-analysis to estimate the extent to which one or more covariates with values defined for each study in the meta-analysis explain any observed heterogeneity in the study effect estimates. When the data are in the form usually found in occupational cohort studies, that is, can be expressed in the form O/E fixed effect and random effects meta-regression analyses may be undertaken even when $O = 0$ in some studies using a Poisson regression or negative binomial regression model respectively.

4.3.5.1 Fixed effect Poisson regression

In a model where the observed number of cases of interest are from an epidemiological study, it is often reasonable to assume an underlying Poisson distribution for the response variable and to describe the impact of explanatory variables, the study characteristics, on their means by some regression function.

To define a basic version of a Poisson regression model, suppose there are observations y_1, y_2, \dots, y_n for the response variable Y_1, Y_2, \dots, Y_n , assumed to be independently distributed Poisson variates with means $\mu_1, \mu_2, \dots, \mu_n$, i.e.

$$f(y_i | \mu_i) = \frac{\mu_i^{y_i}}{y_i!} \exp(-\mu_i).$$

The systematic component of the model is specified by some regression function η , depending on regression parameters β_1, \dots, β_k , with each component relating values x_{i1}, \dots, x_{ik} of explanatory variables that are study-related to respective means, i.e.

$$\mu_i = \eta_i(\beta) = \eta_i(x_{i1}, \dots, x_{ik}; \beta_1, \dots, \beta_k).$$

Often this relationship is such that some monotone transformation g of the means is connected to a linear predictor of explanatory variables,

$$g(\mu_i) = \sum_{j=1}^k x_{ij} \beta_j.$$

In this situation, g is called the link function and the model defined in this manner is an example of a generalised linear model.

For

$$\eta_i(\beta) = \exp\left(\sum_{j=1}^k x_{ij} \beta_j\right)$$

the familiar log-linear model,

$$\log \mu_i = \sum_{j=1}^k x_{ij} \beta_j \quad [79].$$

When dealing with RRs that can be expressed in the form of observed divided by expected, the observed number of cases can be regarded as from the Poisson distribution and the expected number is regarded as fixed (i.e. known without error) in any modelling.

This means that the expected number of cases is regarded as an offset in the model and thus

$$\eta_i(\beta) = \exp\left(\sum_{j=1}^k (x_{ij}\beta_j + offset_i)\right) \text{ and}$$

$$\log \mu_i = \sum_{j=1}^k (x_{ij}\beta_j + offset_i).$$

Thus applying a Poisson regression in a meta-regression involves identifying and coding study characteristics that might be responsible for between study heterogeneity in effect estimates. Provided there are sufficient studies, examining these variables in a Poisson regression may provide some insight into whether there might be reasons due to study characteristics for differences in effect estimates.

4.3.5.2 *Random effects negative binomial regression*

Negative binomial regression is used to model the number of occurrences of an event when the event has extra-Poisson variation, i.e. when there is over-dispersion.

The structure of the model is similar to that for the Poisson regression model, except there is an additional variable for the extra-Poisson variation that follows a gamma distribution with mean 1 and variance α .

Thus

$$Y_i \sim \text{Poisson}(\mu_i^*),$$

where

$$\mu_i^* = \exp\left(\sum_{j=1}^k (x_{ij}\beta_j + offset_i + u_i)\right)$$

and

$$e^{u_i} \sim Gamma(1/\alpha, \alpha),$$

which has expectation 1 and variance α . The term α is referred to as the dispersion parameter and the larger α , the greater the overdispersion or between-study variability.

When $\alpha = 0$, the model corresponds to the Poisson regression model [80].

Note that null models (using the methods described) could be used as additional options to obtaining single pooled estimates of meta-relative risk.

4.4 SIMULATIONS OF META-ANALYSES OF SMRS FROM COHORT STUDIES, WHEN EXPECTED NUMBER OF CASES IS SMALL

As mentioned in 4.3.1.2, the value of k which is used to allow studies with zero relative risk could be 0.5 if based on precedent or alternatively as small as 10^{-8} . This section contains some simulations to illustrate the possible effect that different choices of k might have on the results of a fixed effect meta-analysis.

Meta-analyses of SMRs in which a simple correct solution was available from the original untransformed scale (based on summing of observed and expected cases) were carried out. The number of studies of 30 was chosen because it was similar to the number of effect estimates anticipated to be used in the nasopharyngeal and sinonasal cancer examples.

4.4.1 Basic approach

SMRs were randomly generated using the Poisson random variable to generate a random distribution of observed cases from 30 studies. Thus each study has the potential to have a different number of cases. This was repeated for different underlying true SMRs of 1.0, 1.5 and 2.0. Summary SMRs were generated by summing the observed and expected cases on the unlogged (original) scale. Exact 95 percent confidence intervals (95% CIs) were generated. For comparison fixed effect summary SMRs and 95% CIs on the log-transformed scale were also generated using the ‘metan’ routine in Stata and compared with the estimates derived on the original scale. The results were also analysed using Poisson and negative binomial regression for the case $k = 0$. In addition these calculations were repeated by adding 10 alternative continuity corrections of 10^{-8} , 0.001, 0.01, 0.1 0.25, 0.5, 0.75, 1, 2 and 5 to both the observed and expected cases from each study.

4.4.2 Results when true summary SMR = 1.0

In the following table, there were 12 studies out of 30 that had zero cases randomly generated. The results from the Poisson and negative binomial regression for $k = 0$ were identical at 0.97 (0.67 to 1.39) and the central estimate was confirmed as being the same as the fixed effect summary SMR, but with a slightly different 95% confidence interval.

Table 4.1 - Comparison of results when true summary SMR = 1.0

Fixed effect meta-analysis on original untransformed scale (exact 95% CI)	Continuity Correction k	Fixed effect SMR using Stata routine 'metan' (95% CI) and continuity correction
0.97 (0.65 to 1.39)	0 (i.e. no correction)	2.04 (1.57 to 2.66)
0.97 (0.65 to 1.39)	10^{-8}	2.04 (1.57 to 2.66)
0.97 (0.65 to 1.39)	0.001	2.04 (1.57 to 2.65)
0.97 (0.65 to 1.39)	0.01	2.03 (1.56 to 2.64)
0.97 (0.66 to 1.37)	0.1	1.91 (1.49 to 2.46)
0.97 (0.68 to 1.34)	0.25	1.75 (1.39 to 2.20)
0.98 (0.71 to 1.31)	0.5	1.54 (1.26 to 1.90)
0.98 (0.73 to 1.29)	0.75	1.41 (1.18 to 1.69)
0.98 (0.75 to 1.27)	1	1.32 (1.12 to 1.55)
0.99 (0.79 to 1.22)	2	1.14 (1.02 to 1.28)
0.99 (0.85 to 1.15)	5	1.03 (0.97 to 1.10)

4.4.3 Results when true summary SMR = 1.5

In the following table, there were 6 studies out of 30 that had zero cases randomly generated. The results from Poisson and negative binomial regression were again identical at 1.37 (1.01 to 1.86). As for the case when SMR = 1.0, the summary SMR is that same as that using the meta-SMR, with slightly different 95% confidence intervals.

Table 4.2 - Comparison of results when true summary SMR = 1.5

Fixed effect meta-analysis on original untransformed scale (exact 95% CI)	Continuity Correction k	Fixed effect SMR using Stata routine 'metan' (95% CI) and continuity correction
1.37 (0.98 to 1.85)	0 (i.e. no correction)	2.50 (2.04 to 3.08)
1.37 (0.98 to 1.85)	10^{-8}	2.50 (2.04 to 3.08)
1.37 (0.98 to 1.85)	0.001	2.50 (2.04 to 3.07)
1.36 (0.98 to 1.85)	0.01	2.49 (2.03 to 3.05)
1.33 (0.97 to 1.79)	0.1	2.32 (1.91 to 2.83)
1.29 (0.96 to 1.71)	0.25	2.10 (1.78 to 2.54)
1.24 (0.94 to 1.62)	0.5	1.85 (1.57 to 2.15)
1.20 (0.93 to 1.55)	0.75	1.68 (1.44 to 1.95)
1.18 (0.92 to 1.49)	1	1.55 (1.35 to 1.78)
1.12 (0.91 to 1.36)	2	1.30 (1.17 to 1.44)
1.06 (0.92 to 1.23)	5	1.10 (1.05 to 1.17)

4.4.4 Results when true summary SMR = 2.0

In the following table, there were 4 studies out of 30 that had zero cases randomly generated. The results from the Poisson and negative binomial regression were identical at 1.73 (1.32 to 2.27). As for the other true summary SMR values, the meta-SMR gave the same estimate as the Poisson and negative binomial regression, albeit with slightly different 95% confidence intervals.

Table 4.3 - Comparison of results when true summary SMR = 2

Fixed effect meta-analysis on original untransformed scale (exact 95% CI)	Continuity Correction k	Fixed effect SMR using Stata routine 'metan' (95% CI) and continuity correction
1.73 (1.29 to 2.27)	0 (i.e. no correction)	2.90 (2.45 to 3.43)
1.73 (1.29 to 2.27)	10^{-8}	2.90 (2.45 to 3.43)
1.73 (1.29 to 2.27)	0.001	2.90 (2.45 to 3.43)
1.72 (1.29 to 2.26)	0.01	2.88 (2.44 to 3.40)
1.67 (1.26 to 2.17)	0.1	2.69 (2.29 to 3.17)
1.59 (1.21 to 2.05)	0.25	2.44 (2.09 to 2.85)
1.49 (1.15 to 1.89)	0.5	2.14 (1.86 to 2.46)
1.42 (1.12 to 1.78)	0.75	1.93 (1.70 to 2.19)
1.37 (1.09 to 1.70)	1	1.78 (1.58 to 2.00)
1.24 (1.03 to 1.50)	2	1.45 (1.32 to 1.59)
1.12 (0.97 to 1.29)	5	1.18 (1.12 to 1.24)

4.4.5 Indications for a continuity correction from the simulations

The simulations in the previous section were based on only a single realisation of 30 Poisson random variables. Thus any inferences drawn from them have to be interpreted cautiously. However the following were indicated:

Introduction of k for the analyses on the original scale simply introduces an unwanted bias and is therefore unnecessary.

For $\mu = 1$ exclusion of 12 out of 30 studies that have zero relative risk causes an upward bias in the estimated relative risk such that the risk is doubled and becomes statistically significant at the 5% level of significance. The addition of very small and small constants does not appear to resolve the issue, as $k = 10^{-8}$, 0.001 and 0.01 give very similar estimates of fixed effect meta-SMR. A similar pattern is seen for $\mu = 1.5$ and $\mu = 2$.

For $\mu = 1$, it is only when k gets as large as 5 does the 95% confidence interval for the true meta-SMR include the true value of 1. For $\mu = 1.5$, k was at least 0.75 before the

95% confidence interval included 1.5 and for $\mu = 2$, k was at least 0.5 before 2 was included in the 95% confidence interval.

The optimum value of k depends on the true value of the SMR and it doesn't seem possible to choose it in advance.

These examples are illustrative only, especially with the simulations being based on a single realisation. There doesn't appear to be an optimal value of k , so if an approach is to be adopted the conventional $k = 0.5$ may as well be used.

When k is very small or small, the results of the meta-analyses on the log scale appear to be the same as when $k = 0$. This point can be demonstrated algebraically. As given in Chapter 3, the meta-SMR on the log scale is given by:

$$\overline{\log SMR} = \frac{\sum_{i=1}^n w_i \log SMR_i}{\sum_{i=1}^n w_i}, \text{ where the number of studies is now denoted by } n.$$

$$\text{Here } w_i = \frac{1}{\text{var}(\log SMR_i)}$$

$$\text{and } \text{var}(\log SMR_i) = \frac{1}{O_i}.$$

Thus when k is added to each O and E for each study where $O_i = 0$, we have

$$\overline{\log SMR_i^*} = \frac{\sum_{O_i=0} k \log\left(\frac{k}{E_i + k}\right) + \sum_{O_i \neq 0} O_i \log\left(\frac{O_i}{E_i}\right)}{\sum_{O_i=0} k + \sum_{O_i \neq 0} O_i} = \frac{\sum_{O_i=0} k \log k - \sum_{O_i=0} k \log(E_i + k) + \sum_{O_i \neq 0} O_i \log\left(\frac{O_i}{E_i}\right)}{\sum_{O_i=0} k + \sum_{O_i \neq 0} O_i}.$$

L'Hopital's rule can be used to show that $\lim_{x \rightarrow 0} x \log x \rightarrow 0$.

$$\text{Thus as } k \rightarrow 0, \overline{\log SMR_i^*} \rightarrow \frac{0 - 0 + \sum_{O_i \neq 0} O_i \log\left(\frac{O_i}{E_i}\right)}{0 + \sum_{O_i \neq 0} O_i} = \frac{\sum_{O_i \neq 0} O_i \log\left(\frac{O_i}{E_i}\right)}{\sum_{O_i \neq 0} O_i}$$

which only contains a contribution from those studies where O_i is not equal to zero.

Thus when k is very small, the fixed effect analysis on the log scale which is carried out by adding k to O and E for those studies with zero relative risk is approximately equivalent to excluding those studies from the analysis.

This suggests that adding a constant and taking logs are not reliable ways of processing.

The results from the Poisson and negative binomial regression as give the same summary SMR as the meta-SMR. Indeed both Poisson and negative binomial regression gave the same confidence intervals which were very similar to those obtained from exact Poisson probabilities for the meta-SMR.

As well as simulating multiple realisations, extending with work to include other random effects models would be appropriate.

4.5 SUMMARY OF APPROACHES

The following meta-analyses were undertaken using the formaldehyde and nasopharyngeal cancer and formaldehyde and sinonasal cancer data sets:

- Fixed effect and random effects on the log scale, excluding risk estimates with no observed cases
- Fixed effect and random effects on the log scale, with $k = 0.5$ added to the observed and expected number of cases for each study
- Fixed effect and random effects on the square root scale
- Fixed effect and random effects on the original untransformed scale
- Poisson and negative binomial regression

The analyses were undertaken for all studies combined when studies other than cohort studies were included by generation of pseudo expected numbers of cases and treated in the same way as cohort studies and analyses were undertaken for cohort studies alone.

4.6 COMPARISON OF APPROACHES USING DATA FOR FORMALDEHYDE AND NASOPHARYNGEAL CANCER

The analyses presented in this section include all of the studies included in the analysis in Chapter 5.

Table 4.4 - Headline nasopharyngeal cancer analyses – all studies

Analysis	No. of estimates	Meta-RR	SE (Meta-RR)	95% CI	P-value for heterogeneity
Log scale					
FE, excluding studies with O = 0	12	1.38	0.15	1.11 to 1.71	0.137
RE, excluding studies with O = 0	12	1.41	0.21	1.06 to 1.88	0.137
FE, adding k = 0.5 to O & E for studies with O = 0	16	1.31	0.14	1.07 to 1.62	0.867
RE, adding k = 0.5 to O and E for studies with O = 0	16	1.31	0.14	1.07 to 1.62	0.867
Square root scale					
FE	16	1.19	0.02	1.15 to 1.23	<0.001
RE	16	1.05	0.14	0.80 to 1.33	<0.001
Original scale					
FE	16	1.19	0.13	0.94 to 1.44	<0.001
RE	16	1.41	0.41	0.60 to 2.22	<0.001
Poisson regression	16	1.23	0.08	1.08 to 1.40	0.039
Negative binomial regression	16	1.26	0.13	1.02 to 1.56	0.039
Abbreviations: FE = Fixed effect RE = Random effects Notes: All non-cohort studies are included in the square root and regression analyses by assuming E = O / RR for each study. For the original scale analyses for cohort studies, it was assumed that SE(RR) = sqrt(E), and for non-cohort studies, SE(RR) was generated from SE(log RR) using the delta method (see section 4.3.2).					

The naïve analysis on the logarithmic scale in Table 4.4 excluded 4 studies and gave statistically significantly raised meta-RRs from both the fixed effect and random effects analyses. On this scale, the between-study heterogeneity was not statistically significant. Adding $k = 0.5$ to the observed and expected values for each study did not appreciably alter the findings.

For the analysis on the square root scale, the fixed effect analysis gave a statistically significantly raised meta-RR, but in the presence of significant heterogeneity on this scale, the random effect meta-RR is preferred and was not statistically significantly raised.

The analysis on the original scale, however, was not statistically significant raised in the fixed effect analysis and in the presence of statistically significant heterogeneity, was not significantly raised in the random effects analysis either. Interestingly, although the Poisson and negative binomial regression effect estimates give similar meta-RRs, these were both marginally statistically significantly raised.

Table 4.5 repeated the same analyses, but restricted to just the cohort studies. On the logarithmic scale (excluding the studies with zero observed cases), the fixed effect meta-RR was significantly raised in the presence of significant between-study heterogeneity; however the random effects analysis was not significantly raised.

Adding $k = 0.5$ to O and E gave analyses that were not statistically significantly raised from either the fixed effect or random effects analyses.

Neither the fixed effect or random effects analyses on the square root scale gave meta-RRs that were raised, nor did the weighted average on the original scale or regression analyses.

Table 4.5 - Nasopharyngeal cancer analyses – cohort studies only

Analysis	No. of estimates	Meta-RR	SE (Meta-RR)	95% CI	P-value for heterogeneity
Log scale					
FE, excluding studies with O = 0	5	1.75	0.45	1.06 to 2.89	0.027
RE, excluding studies with O = 0	5	2.40	1.22	0.89 to 6.50	0.027
FE, adding k = 0.5 to O & E for studies with O = 0	9	1.39	0.29	0.93 to 2.08	0.715
RE, adding k = 0.5 to O and E for studies with O = 0	9	1.39	0.29	0.93 to 2.08	0.715
Square root scale					
FE	9	0.98	0.06	0.85 to 1.10	<0.001
RE	9	0.99	0.45	0.29 to 2.06	<0.001
Original scale					
FE	9	1.27	0.26	0.76 to 1.78	<0.001
RE	9	2.31	1.15	0.06 to 4.57	<0.001
Poisson regression	9	1.27	0.29	0.81 to 1.99	0.479
Negative binomial regression	9	1.27	0.30	0.80 to 2.01	0.479
Abbreviations: FE = Fixed effect RE = Random effects Notes: All non-cohort studies are included in the square root and regression analyses by assuming $E = O / RR$ for each study. For the original scale analyses for cohort studies, it was assumed that $SE(RR) = \sqrt{E}$, and for non-cohort studies, $SE(RR)$ was generated from $SE(\log RR)$ using the delta method (see section 4.3.2).					

4.7 COMPARISON OF APPROACHES USING DATA FOR FORMALDEHYDE AND SINONASAL CANCER

Table 4.6 - Headline sinonasal cancer analyses – all studies

Analysis	No. of estimates	Meta-RR	SE (Meta-RR)	95% CI	P-value for heterogeneity
Log scale					
FE, excluding studies with O = 0	20	1.31	0.10	1.13 to 1.52	0.009
RE, excluding studies with O = 0	20	1.34	0.16	1.06 to 1.69	0.009
FE, adding k = 0.5 to O & E for studies with O = 0	28	1.25	0.09	1.08 to 1.45	0.207
RE, adding k = 0.5 to O and E for studies with O = 0	28	1.25	0.11	1.05 to 1.50	0.207
Square root scale					
FE	28	1.22	0.02	1.18 to 1.25	<0.001
RE	28	0.89	0.13	0.66 to 1.16	<0.001
Original scale					
FE	28	1.01	0.08	0.85 to 1.17	0.388
RE	28	1.01	0.09	0.83 to 1.18	0.388
Poisson regression	28	1.32	0.07	1.19 to 1.47	<0.001
Negative binomial regression	28	1.29	0.17	1.00 to 1.67	<0.001
Abbreviations: FE = Fixed effect RE = Random effects Notes: All non-cohort studies are included in the square root and regression analyses by assuming $E = O / RR$ for each study. For the original scale analyses for cohort studies, it was assumed that $SE(RR) = \sqrt{E}$, and for non-cohort studies, $SE(RR)$ was generated from $SE(\log RR)$ using the delta method (see section 4.3.2).					

The naïve analysis for sinonasal cancer on the logarithmic scale (Table 4.6), excluded eight effect estimates and was statistically significantly raised for both the fixed effect and random effects analysis. As for the nasopharyngeal cancer analyses, adding $k = 0.5$ to the observed and expected values for each study did not materially alter the findings from these analyses. The meta-RR from the fixed effect analysis on the square root scale was statistically significantly raised, but in the presence of highly statistically significant between-study heterogeneity, was not significantly raised in the random effects analysis. The meta-RR from both the fixed effect and random effects analysis on the original scale were not statistically significantly raised, but as for the nasopharyngeal cancer findings, were both marginally statistically significantly raised from the Poisson and negative binomial regression analyses. Restricting the sinonasal cancer analyses to cohort studies only (Table 4.7) did not produce any statistically significant findings, regardless of the scale or whether the analysis was fixed effect or random effects.

Table 4.7 - Sinonasal cancer analyses – cohort studies only

Analysis	No. of estimates	Meta-RR	SE (Meta-RR)	95% CI	P-value for heterogeneity
Log scale					
FE, excluding studies with O = 0	3	1.06	0.26	0.65 to 1.70	0.835
RE, excluding studies with O = 0	3	1.06	0.26	0.65 to 1.70	0.835
FE, adding $k = 0.5$ to O & E for studies with O = 0	10	0.94	0.19	0.64 to 1.39	0.981
RE, adding $k = 0.5$ to O and E for studies with O = 0	10	0.94	0.19	0.64 to 1.39	0.981

Square root scale					
FE	10	0.72	0.04	0.64 to 0.71	<0.001
RE	10	0.10	0.10	0.00 to 0.40	<0.001
Original scale					
FE	10	0.86	0.20	0.47 to 1.26	0.914
RE	10	0.86	0.20	0.47 to 1.26	0.914
Poisson regression	10	0.86	0.19	0.56 to 1.32	1.000
Negative binomial regression	10	0.86	0.19	0.56 to 1.32	1.000
Abbreviations: FE = Fixed effect RE = Random effects Notes: All non-cohort studies are included in the square root and regression analyses by assuming $E = O / RR$ for each study. For the original scale analyses for cohort studies, it was assumed that $SE(RR) = \sqrt{E}$, and for non-cohort studies, $SE(RR)$ was generated from $SE(\log RR)$ using the delta method (see section 4.3.2).					

When the analyses are restricted to the cohort studies only, the fixed effect analysis on the original scale is not statistically significant and there is no significant heterogeneity. None of the results on the log or square root scales provide any statistically significant mean relative risks.

4.8 DISCUSSION AND CONCLUSIONS

This chapter began with a review of the approaches that have been employed in the literature to date for meta-analyses of rare health outcomes. The methods produced have essentially been in the context of the meta-analyses of controlled clinical trials. No clear conclusions are drawn about which is the optimum approach and none of the approaches deal with meta-analyses in occupational epidemiology.

For meta-analyses in occupational epidemiology three methods have been examined in the context of meta-analyses of occupational exposure to formaldehyde and risks of nasopharyngeal cancer and sinonasal cancer. Cohort studies have been examined in isolation as well as in combination with other study designs.

When the true SMR equals one, $k = 5$ provided the least biased effect estimate when the analysis was carried out on the logarithmic scale. Clearly it does not make sense to use such a large value of k in practice. When the true SMR equals 1.5, the optimum value of k appears to be $k = 1$ and for SMR = 2, $k = 0.75$ appears to give the least biased estimate of effect. Given that these simulations were based on a single realization it is difficult to generalize the findings. However, what is clear is that analysis on the logarithmic scale is not appropriate for rare diseases, especially when plausible alternative analytical approaches exist on the original untransformed scale. The findings for both nasopharyngeal cancer and sinonasal cancer both gave meta-RRs that were statistically significantly raised when the analysis was carried out on the logarithmic scale. However, analyses on the original scale using 'metan' in Stata or using Poisson or negative binomial regression suggest that the findings of an excess meta-RR are not robust and that determination of a causal association is not yet strongly supported.

Analyses on the log scale causing the exclusion of studies with no observed cases are clearly biased and when the disease is rare has significant potential to produce misleading results. Analyses on the original scale or Poisson and negative binomial regression are to be preferred. Any meta-analysis of a rare disease should also include a good sensitivity analysis as a check on the robustness of any findings and therefore before any inferences are drawn.

5 REVIEW AND META-ANALYSIS OF THE EPIDEMIOLOGICAL EVIDENCE FOR THE CARCINOGENICITY OF FORMALDEHYDE IN RELATION TO LUNG CANCER, NASOPHARYNGEAL CANCER AND SINONASAL CANCER

5.1 INTRODUCTION

An earlier analysis of the National Cancer Institute (NCI) cohort [82] suggested that the association between nasopharyngeal cancer and formaldehyde exposure might be causal. An assessment of robustness with respect to alternative methods of data analysis, including categorisations of formaldehyde exposure, was subsequently undertaken and concluded that the nasopharyngeal excess was centred on one plant, which had been separately analysed and no excess found [83], and that overall there was considerable uncertainty over whether or not the association was causal [84]. The NCI cohort analysis probably prompted the International Agency for Research on Cancer (IARC) to update their evaluation of formaldehyde and resulted in a determination for the first time of the carcinogenicity of formaldehyde in relation to nasopharyngeal cancer [85]. The IARC evaluation was apparently heavily dependent on the NCI cohort, which had previously been criticized as not being robust and this prompted Marsh *et al* to look again at the NCI cohort, this time looking at the model specification and the degree of instability in the risk estimates. Marsh *et al* reiterated their earlier claim that the NCI cohort did not support the determination of a causal association for occupational exposure to formaldehyde and nasopharyngeal cancer [86]. Marsh *et al* subsequently carried out further analyses to determine whether the association at the plant with the large nasopharyngeal excess could have been due to factors other than formaldehyde exposure and concluded that the association might be due to silver

smithing and other metal work [86]. It was subsequently admitted by the NCI that there may have been incomplete follow-up of their cohort which may have affected the subsequent re-analyses by Marsh *et al* [87].

The most recent of IARC evaluation was published in 2012 [88]. In relation to the evidence for carcinogenicity in humans, IARC concluded that there was sufficient evidence for the carcinogenicity in relation to nasopharyngeal cancer, that a positive association was found for sinonasal cancer, and that several studies showed statistically significant positive associations for lung cancer. This and a further follow-up of the NCI cohort [89], prompted Marsh *et al* to reiterate all of their earlier concerns [90]. They subsequently published an updated re-analysis and again reiterated their earlier concerns about the analysis and interpretation of the NCI cohort [91].

This series of publications highlights the question as to what evidence exists, aside from the NCI cohort, for the association between occupational exposure to formaldehyde and the risk of nasopharyngeal. Further examination of the epidemiological evidence in relation to sinonasal cancer and lung cancer is also warranted. For the two rarer cancers, the impact of the exclusion of studies with no exposed cases on the totality of the evidence is also examined for its potential impact on any inferences drawn.

5.1.1 Distribution of absorption of formaldehyde in the human body

Formaldehyde is produced endogenously in the body and is an essential metabolic intermediate in all human cells [92]. Formaldehyde is also a skin irritant and can cause allergic contact dermatitis [92]. The production of DNA-protein cross-links induced by a single six hour exposure to formaldehyde has been compared in several regions of the upper respiratory tract of F-344 rats and rhesus monkeys. Concentrations of the cross-links in the nasal mucosa were significantly lower in the monkeys than in the rats. Cross-links were also detected in the nasopharynx and trachea of the monkeys, but not

in the nasal sinus, proximal lung or the bone marrow [93]. It is possible, however, for formaldehyde to reach the lung when adsorbed to particulates [92]. The distance from the nostrils to the nasopharynx is usually proportional to head and snout length, which varies widely in rodents, in which atrioturbinates in the nasal vestibule, a structure which is lacking in humans, act as baffles which deflect large volumes of air away from the nasopharynx [94]. The site-specific flux of formaldehyde into tissue is heavily dependent on the proportion and concentration of the inhaled formaldehyde coming into contact with the tissue [95]. Thus, it is reasonable to conclude that any carcinogenic risk in humans might extend to respiratory organs beyond the nasal cavity, other than the nasal sinuses, but that the exact dose to a particular tissue depends on a number of factors including the level of formaldehyde inhaled. There is also experimental evidence that formaldehyde causes nasal cancer in rats [96].

5.1.2 The known and suspected risk factors for nasopharyngeal cancer, sinonasal cancer and lung cancer

Cancer of the nasopharynx is rare in England and Wales, with about 200 new cases per year in adults. Annual incidence rates are about 0.4 per 100,000 in men and 0.2 per 100,000 in women [97]. Nasopharyngeal cancer is believed to result from a combination of genetic susceptibility, infection with Epstein-Barr virus and regular consumption of salted fish in childhood [98]. Smoking is not regarded as a risk factor for nasopharyngeal cancer.

Cancers of the nasal cavity and paranasal sinuses are uncommon. There are about 400 new cases per year in England and Wales, with incidence rates of 1 per 100,000 in men and 0.5 per 100,000 in women [97]. Wood dust and selected nickel compounds, especially oxides and sulphides involved in nickel refining cause cancer of the nasal cavities and paranasal sinuses. Hexavalent chromium compounds and untreated and

mildly treated mineral oils cause cancer of the nasal sinuses according to the IARC classification [99]. Smoking is not regarded as a risk factor for cancers of the nasal cavity or paranasal sinuses.

Lung cancer is the most common cancer in men in England and Wales, accounting for about one in four cancers (25,000 cases per year), and the third most common in women (1 in 9; 12,000 cases per year). Lung cancer accounts for 6% of deaths from all causes. Incidence rates are typically for developed countries, about 100 per 100,000 per year for men and 45 per 100,000 for women [97]. Aside from tobacco smoking, IARC has classified many exposures as definitely causing lung cancer: ionising radiation including X-rays, gamma-rays, neutrons and radon gas; asbestos; crystalline silica; talc containing asbestiform fibres; arsenic and arsenical compounds; beryllium; cadmium and cadmium compounds; selected nickel compounds including combinations of oxides and sulphides involved in nickel refining; coal and tar pitches; untreated and mildly treated mineral oils; soots; bis(chloromethyl) ether and technical grade chloromethyl ether; 2,3,7,8-tetrachlorodibenzo para-dioxin; passive smoking; mustard gas; and strong acid mists containing sulphuric acid [99]. Also classified as probably lung carcinogens are: benz[a]anthracene; benz[a]pyrene; dibenz[a,h]anthracene; diesel engine exhaust; alpha-chlorinated toluenes; epichlorohydrin; and nonarsenical insecticides [99].

5.1.3 Formaldehyde and its industrial use

Formaldehyde has the molecular formula CH_2O . At room temperature it is a colourless gas with a pungent odour. In 1992, about 12 million tonnes were produced worldwide. Formaldehyde is produced by catalytic vapour phase oxidation of methanol, and is most commonly commercially available as an aqueous solution commonly referred to as 'formalin'. The widest use of formaldehyde is in the production of resins with urea,

phenol and melamine and, to a small extent, their derivatives. Formaldehyde-based resins are used in adhesives and impregnating resins in the manufacture of particle-board, plywood, furniture and other wood products. The resins are also used in the production of curable moulding materials and as raw materials for surface coatings and controlled-release nitrogen fertilisers. They are also used in the textile, leather, rubber and cement industries. Further uses are as binders for foundry sand, stonewool and glasswool mats in insulating materials, abrasive paper and brake linings. Formaldehyde itself is used for preservation and for disinfection, for example it is used in human and veterinary drugs and biological materials, for disinfecting hospital wards and for preserving and embalming biological specimens. It is used as an antimicrobial agent in many cosmetic products. It is also used directly to inhibit corrosion, in mirror finishing, electroplating, in the electrodisposition of printed circuits, and in photographic film development [92]. Formaldehyde is also present in tobacco smoke [100].

5.1.4 Formaldehyde exposure in different occupational circumstances

This section contains a brief overview of the formaldehyde exposures encountered in a range of occupational circumstances that have been the subject of epidemiological investigations. Table 5.1 sets out the results of a selection of the results of hygiene surveys that have been carried out in relation to exposure to formaldehyde in the workplace. It is presented for illustration purposes only and is not meant to be a full review, but rather a brief overview.

Table 5.1 - Brief overview of formaldehyde exposures in workplaces

Occupational circumstances	Exposure range (ppm)	Arithmetic mean exposure level (ppm)	Personal (P) or Area (A) sample	Reference
Formaldehyde manufacture	ND ⁽¹⁾ -4.78		P	Stewart <i>et al</i> , 1987[101]
Funeral homes	0.09-5.26 0.12-5.64 0.11-4.12 0.05-8.37 0.18-0.43 0.00-2.11 0.31-8.72 0.23-7.52 0.28-8.15	0.25-1.39 (range across 6 funeral homes) 0.78 0.80 0.61 0.27 0.88 2.58 2.03 2.16	A P P A P P P A P	Kerfoot & Mooney, 1975 [102] Korczynski, 1994 [103] Williams <i>et al</i> , 1984[104] Stewart <i>et al</i> , 1992 [105]
Anatomy laboratories	0.30-2.63 0.07-2.94 0.24-5.87 0.31-6.77 0.18-1.29	1.25 ⁽²⁾ 1.24 1.69 1.53 0.50	P P P P A	Skisak, 1983 [106] Akbar-Khanzedah <i>et al</i> , 1994 [107] Perkins & Kimbrough, 1985 [108]
Pathology laboratories	<0.10-13.57 2.2-7.9		P A	Coldiron <i>et al</i> , 1983[109] Bernstein <i>et al</i> , 1984[110]
Garment factory	0.46-1.13 0.43-1.30 0.42-1.40	0.82 0.92 1.05	P A A	Luker & Van Houten, 1990 [111]

Occupational circumstances	Exposure range (ppm)	Arithmetic mean exposure level (ppm)	Personal (P) or Area (A) sample	Reference
Fertilizer production	0.2-1.9		A	Bernstein <i>et al</i> , 1984 [110]
Dyestuffs manufacture	<0.1-5.8		A	Bernstein <i>et al</i> , 1984 [110]
Resins manufacture	<0.1-5.5 ND-43.2		A P	Bernstein <i>et al</i> , 1984 [110] Stewart <i>et al</i> , 1987[101]
Iron foundry	0.12-0.8		A	Bernstein <i>et al</i> , 1984 [110]
Treated paper	0.14-0.99		A	Bernstein <i>et al</i> , 1984 [110]
Plywood manufacture	1.0-2.5		A	Bernstein <i>et al</i> , 1984 [110]
Furniture manufacture	<0.1-6.1	1.16 0.06	A P	Priha <i>et al</i> , 1986 [112] Sass-Kortsak <i>et al</i> , 1986 [113]
Woodworking		0.10	P	Partanen <i>et al</i> , 1985 [114]
Urea-formaldehyde foam application	<0.08-2.4		A	Bernstein <i>et al</i> , 1984[110]
Molding compound manufacture	ND-79.5		P	Stewart <i>et al</i> , 1987 [101]
Molded plastic products manufacture	ND-44.5		P	Stewart <i>et al</i> , 1987 [101]

Occupational circumstances	Exposure range (ppm)	Arithmetic mean exposure level (ppm)	Personal (P) or Area (A) sample	Reference
Decorative laminates manufacture	ND-13.46		P	Stewart <i>et al</i> , 1987 [101]
Plywood panelling manufacture	ND-2.01		P	Stewart <i>et al</i> , 1987 [101]
Photographic film manufacture	ND-4.49		P	Stewart <i>et al</i> , 1987 [101]
(1) ND = Not detectable (2) Estimated from presented summary of data				

Since non-occupational exposure to formaldehyde is ubiquitous, all work contributes to total human exposure. The mean concentrations of formaldehyde measured in the 1980s during the manufacture of formaldehyde and formaldehyde-based resins were below 1 ppm. In the wood and pulp and paper industries, the highest mean concentrations are usually in gluing departments and the mean levels were usually >1 ppm before the mid-1970s, but have been below that since. The levels of formaldehyde measured in particle-board mills before the mid-1970s were high – often well over 2 ppm – but the development of glues with lower formaldehyde content and improved ventilation have decreased the levels to about 1 ppm or below. Measurements of formaldehyde in the garment industry from the 1980s were relatively low, averaging 0.1-0.2 ppm. Phenol-formaldehyde resins are commonly used to bind man-made mineral fibre products; measurements in the 1980s showed mean concentrations of 0.1-0.2 ppm of formaldehyde. The mean concentrations of formaldehyde measured during the coating of photographic film and during the development of photographs are usually well below 1 ppm. The concentration of formaldehyde in the air during embalming depends on the content of the embalming fluid, type of the body, ventilation and work practices; the mean level is about 1 ppm. The mean concentration when used in hospitals for disinfection ranges from 0.1 to 0.8 ppm. The mean formaldehyde concentration in histopathology laboratories is usually around 0.5 ppm [92].

5.1.5 Brief Summary of the toxicological evidence

Formaldehyde has acute effects in mammals. Inhalation of high concentrations ($> 120 \text{ mg/m}^3$) of formaldehyde causes hypersalivation, acute dyspnea, vomiting, muscular spasms, convulsions and finally death. Histopathological examination shows respiratory tract irritation, bronchioalviolar constriction and lung oedema.

Formaldehyde is irritating to the eyes, and aqueous solutions of formaldehyde (0.1% to 20%) are irritating to the skin of rabbits [115].

In humans transient and reversible sensory irritation of the eyes and respiratory tract has been observed in clinical studies and epidemiological surveys. The odour threshold for most people ranges between 0.5 and 1 ppm. In general, eye irritation, the most sensitive endpoint, is associated with airborne concentrations beginning in the range 0.3 to 0.5 ppm. Formaldehyde causes skin irritation and has corrosive properties when ingested. Contact dermatitis can occur at concentrations as low as 30 ppm [115].

Formaldehyde is a highly reactive gas that is absorbed quickly at the point of contact and is also produced by endogenous metabolism. It is rapidly metabolised such that exposure to high concentrations (up to 15 ppm in rats) does not result in increased blood concentrations. Repeated formaldehyde exposure causes toxic effects only in the tissues of direct contact after inhalation, and oral or dermal exposure is characterised by local cytotoxic destruction and subsequent repair of the damage. The typical locations of lesions in experimental animals were the nose after inhalation, the stomach after oral administration and the skin after dermal application. The nature of the lesions depends on the inherent abilities of the tissues involved to respond to the noxious event and the local concentration of the substance [115].

Formaldehyde is weakly genotoxic and is able to induce chromosomal aberrations in mammalian cells. DNA-protein crosslinks are a sensitive measure of DNA modification by formaldehyde. However, the genotoxic effect is limited to those cells, which are in direct contact with formaldehyde. Formaldehyde is therefore a direct acting locally effective mutagen [115].

Chronic inhalation of concentrations of 10 ppm and higher can lead to clear increases in nasal tumour incidence in rats. Most of the nasal tumours are squamous cell carcinomas. Marked non-neoplastic pathological lesions of the nasal epithelium accompanied them. No increased incidence has been found in other organs after

inhalation and via administration routes other than inhalation does not result in local or systemic tumour formation. The damage of nasal tissue plays a crucial role in the tumour induction process, since nasal cancer is only found at concentrations inducing epithelial degeneration and increased cell proliferation. Thus the stimulation of cell proliferation seems to be an important pre-requisite for tumour development. Although formaldehyde exhibits some genotoxic activity, the correlation between cytotoxicity, cell proliferation and the induction of nasal cancer in rats provides a convincing scientific basis for the aetiology of the carcinogenic response to be cytotoxicity driven. In contrast, no significant numbers of tumours have been seen in mice and Syrian hamsters following chronic exposures to concentrations up to 14.3 or 30 ppm, respectively. These clear species differences appear to be related, in part, to the local dosimetry and disposition of formaldehyde in nasal tissues. Species differences in nasal anatomy and respiratory physiology may have a profound effect on susceptibility to formaldehyde induced nasal tumours [115].

5.1.6 IARC reviews of the carcinogenicity of formaldehyde

In their 1995 review, IARC classified formaldehyde as probably carcinogenic to humans [92]. A more recent review by the IARC working group resulted in a classification of formaldehyde as definitely carcinogenic to humans [116]. The latest IARC monograph noted that one industrial cohort study had both a strong overall association between exposure to formaldehyde and nasopharyngeal cancer and the most elevated risks were in the highest exposure category. In addition, however, there were also positive associations in many of the case-control studies, particularly the larger ones and those with higher-quality exposure assessment. Although there were no associations in the two other large industrial cohort studies, the expected number of cases in both of those studies was quite small. Thus IARC concluded, given it was unlikely that bias or confounding could explain the observed association, that

occupational exposure to formaldehyde causes nasopharyngeal cancer in humans.

IARC also noted that many case-control studies showed positive associations for exposure to formaldehyde and sinonasal cancer, some with evidence of an exposure-response pattern. However, many of the studies included co-exposure to wood dust, which is strongly associated with sinonasal cancer. The industrial cohorts show no such association and this could be due to a lack of statistical power, or could indicate that uncontrolled confounding to wood dust partially explains the observed associations in case-control studies. Finally in relation to lung cancer IARC noted that several studies identified statistically significant positive associations with formaldehyde exposure, but that the results were inconsistent.

5.1.7 Previous meta-analyses of formaldehyde and respiratory cancer

To date, eight meta-analyses have been published examining the carcinogenicity of the respiratory system from workplace exposures to formaldehyde (see Table 5.2 for key findings) [76, 117-123].

The first of these examined deaths among males for a variety of cancer sites, including the buccal cavity and pharynx, the nose and the lung from nine different studies [117]. Analysis of mortality among workers at chemical plants gave rise to a meta-RR for buccal cavity and pharynx cancer of 0.74 (0.24 to 1.74), based on 5 deaths from three studies. There were no deaths from nasal cancer against 1.07 expected and the meta-RR for lung cancer was 0.95 (0.83 to 1.09) based on 216 deaths from three studies. For exposure among male pathologists, anatomists and morticians, the meta-RR for buccal cavity and pharynx cancer was 0.81 (0.48 to 1.29), based on 18 deaths from four studies. There were no nasal cancer deaths against 1.8 expected, and the meta-RR for lung cancer was 0.80 (0.68 to 0.94) based on 162 deaths from six studies, and this was a statistically significant deficit. The authors reported that there had been two case-

control studies of formaldehyde that had detected no increased risk for lung cancer.

They concluded that the data provided little evidence for the carcinogenicity of formaldehyde in humans at site of contact, for example the buccal cavity and pharynx and the nose, and that the number of deaths from lung cancer was clearly not in excess among formaldehyde-exposed groups [117].

Table 5.2 – Results of previous meta-analyses of formaldehyde in relation to nasopharyngeal cancer, sinonasal cancer and lung cancer

Meta-analysis number	Meta-analysis/Cancer Site	Industrial Cohorts Meta-RR (95% CI) [No. studies, No. cases]	Cohorts of Professionals (Anatomists, Pathologists and Morticians) Meta-RR (95% CI) [No. studies, No. cases]	Non-nested case-control Studies Meta-RR (95% CI) No. studies/No. exposed cases
1	Levine <i>et al</i> , 1984 [117] Buccal cavity & pharynx Nose Lung	 0.74 (0.24 to 1.74) [3, 5] 0.00 (0.00 to 2.80) [3, 0] 0.95 (0.83 to 1.09) [3, 216]	 0.81 (0.48 to 1.29) [4, 18] 0.00 (0.00 to 1.66) [5, 0] 0.80 (0.68 to 0.94) [6, 162]	 Two case-control studies detected no increased risk for lung cancer
2	Nelson <i>et al</i> , 1986 [118] Buccal cavity & pharynx Nose Lung	 1.64 (0.79 to 3.01) [3, 10] 0.00 (0.00 to 3.35) [6, 0] 1.09 (0.96 to 1.24) [6, 245]	 1.04 (0.55 to 1.35) [4, 21] 0.00 (0.00 to 2.05) [5, 0] 0.74 (0.64 to 0.86) [5, 178]	 Three case-control studies provide indications of a possible association for nasal cancer. Two case-control studies detected no increased risk for lung cancer

Meta-analysis number	Meta-analysis/Cancer Site	Industrial Cohorts Meta-RR (95% CI) [No. studies, No. cases]	Cohorts of Professionals (Anatomists, Pathologists and Morticians) Meta-RR (95% CI) [No. studies, No. cases]	Non-nested case-control Studies Meta-RR (95% CI) No. studies/No. exposed cases
3	Blair <i>et al</i> , 1990 [119](1) Nasopharynx Nose Lung	1.22 (0.86 to 1.73) [4, 36] 1.07 (0.81 to 1.38) [12, 60] 1.06 (1.02 to 1.14) [14, 1181]	2.22 (0.61 to 5.69) [2, 4] 0.42 (0.01 to 2.32) [4, 1] 0.89 (0.80 to 0.95) [13, 511]	Results presented for all study types combined
4	Partanen, 1993 [124] Nasopharynx Sinonasal Lung	1.74 (1.21 to 2.41) [4, 36] 1.19 (0.96 to 1.46) [11, 93] 1.11 (1.03 to 1.19) [14, 833]	2.22 (0.61 to 5.69) [2, 4] 0.42 (0.01 to 2.32) [4, 1] 0.88 (0.80 to 0.95) [13, 511]	Results presented for all study types combined
5	Collins <i>et al</i> , 1997 [76] Nasopharynx Sinonasal Lung	1.36 (0.50 to 2.97) [6, 6] 0.6 (0.1 to 1.7) [6, 3] 1.1 (1.0 to 1.2) [6, 757]	0.91 (0.25 to 2.33) [8, 4] 0.00 (0.0 to 1.0) [6, 0] 0.77 (0.71 to 0.83) [8, 562]	1.3 (0.9 to 2.1) [6, 445] 1.8 (1.4 to 2.3) [11, 933] 0.8 (0.7 to 1.0) [6, 301]
6	Bosetti <i>et al</i> , 2008 [121] Nasopharynx Sinonasal Lung	1.33 (0.61 to 2.53) [3, 9] 1.01 (0.33 to 2.35) [3, 5] 1.07 (1.01 to 1.12) [6, 1459]	Not presented 0.00 (0.00 to 3.51) [7, 0] 0.68 (0.63 to 0.74) [13, 703]	Only considered cohort studies.

Meta-analysis number	Meta-analysis/Cancer Site	Industrial Cohorts Meta-RR (95% CI) [No. studies, No. cases]	Cohorts of Professionals (Anatomists, Pathologists and Morticians) Meta-RR (95% CI) [No. studies, No. cases]	Non-nested case-control Studies Meta-RR (95% CI) No. studies/No. exposed cases
7	Bachand <i>et al</i> , 2010 [125] Nasopharynx (No data presented for sinonasal cancer or for lung cancer)	0.72 (0.40 to 1.29) [11, 39]	Not examined.	1.22 (1.00 to 1.50) [7, 502]
8	Binazzi <i>et al</i> , 2015 [123] Sinonasal cancer (No data presented for nasopharyngeal cancer or for lung cancer)	1.09 (0.66 to 1.79) [2, not presented]	Not separately examined.	1.68 (1.37 to 2.06) [6, not presented]
	Notes: (1) For Blair <i>et al</i> , studies of non-occupational exposures were grouped with exposures to professionals. In addition, where confidence intervals were not presented, they have been calculated using exact methods.			

The second meta-analysis also examined mortality for a variety of cancer sites including cancers of the buccal cavity and pharynx, the nose and the lung from 12 different studies [118]. Among studies of chemical and garment industry workers, the meta-RR for cancer of the buccal cavity and pharynx was 1.64 (0.79 to 3.01) based on 10 deaths from three studies. There were no deaths from nasal cancer, with 1.1 expected, and the meta-RR for lung cancer was 1.09 (0.96 to 1.24) based on 245 deaths from six studies. Among studies of pathologists, anatomists and morticians, the meta-RR for cancer of the buccal cavity and pharynx was 1.04 (0.55 to 1.35) based on 21 deaths from four studies. There were no deaths from nasal cancer against 1.8 expected, and the meta-RR

for lung cancer was 0.74 (0.64 to 0.86), based on 178 deaths from five studies, a statistically significant deficit. As in the report of the first meta-analysis [117], there were two case-control studies of formaldehyde exposure that detected no increased risk for lung cancer. However, there were also three other case-control studies that provided indications of a possible association between formaldehyde exposure and nasal cancer. On the basis of no deaths from nasal cancer in SMR or PMR studies, but three case-control studies suggesting the possibility of an increased risk of nasal cancer, the authors concluded that there was limited epidemiological evidence for the human carcinogenicity of formaldehyde [118].

The third meta-analysis also examined mortality for a variety of cancers in over 30 discrete study populations [119]. Among industrial worker cohorts the meta-RR for nasopharyngeal cancer was 1.22 (0.86 to 1.73), based on 36 cases from four studies. The meta-RR for nasal cancer was 1.07 (0.81 to 1.38) based on 60 deaths from 11 studies, and the meta-RR for lung cancer was 1.06 (1.02 to 1.14) based on 1181 deaths from 14 studies, a statistically significant excess. Among professional and non-occupationally exposed groups, the meta-RR for nasopharyngeal cancer was 2.22 (0.61 to 5.69) based on four deaths from two studies. For nasal cancer the meta-RR was 0.42 (0.01 to 2.32), based on a single death from four studies. The meta-RR for lung cancer was 0.88 (0.80 to 0.95) based on 511 deaths from 13 studies, a statistically significant deficit. The risk of nasal cancer was evaluated by exposure level or duration of exposure to formaldehyde. When the observed and expected numbers from the various studies were combined, no exposure response gradient was evident from the combined RR values. In contrast, the data for nasopharyngeal cancer showed a statistically significantly positive trend of increased mortality risk with increased exposure. There were no significant trends between level of formaldehyde exposure and risk of death from lung cancer. The authors concluded that it was likely that the excesses of

nasopharyngeal cancer observed were caused by exposure to formaldehyde, and that an association with nasal cancer was plausible, but somewhat less persuasive than that for nasopharyngeal cancer. They also concluded that although a role for formaldehyde in the excess of lung cancer could not be dismissed, inconsistencies among and within studies of industrial workers suggested that the association was not causal. They went on to suggest that other interpretations were possible and that formaldehyde might be an effective carcinogen only in the presence of other exposures which were not consistent from study to study, or it may be a weak carcinogen whose effect was easily masked by the presence of lung carcinogens that varied from study to study [119].

In order to further clarify the findings of the third meta-analysis, a fourth meta-analysis has been published [124] which scrutinised the findings of the third meta-analysis [119] in relation to respiratory system cancers. The main difference between the third meta-analysis [119] and the fourth meta-analysis [124] lay in the choice of the risk estimates used in the meta-analysis. Despite a fair number of changes in the input values, the results of the reanalysis agreed generally well with those of the original analysis. Thus the author concluded that cancers of the nasopharynx and nose appear to be the most likely targets for the carcinogenic action of formaldehyde in humans that it remained unlikely that workplace exposures to formaldehyde posed any substantial cancer hazard among humans [124].

A fifth meta-analysis of formaldehyde exposure and respiratory system cancers included results from 47 studies [76]. In industrial cohorts, the meta-RR for nasopharyngeal cancer was 1.36 (0.50 to 2.97), based on six deaths from six studies. The meta-RR for sinonasal cancer was 0.6 (0.1 to 1.7), based on three deaths from six studies. For lung cancer the meta-RR was 1.1 (1.0 to 1.2), based on 757 deaths from six studies, a result that is marginally statistically significant. Among cohorts of pathologists and embalmers, the meta-RR for nasopharyngeal cancer was 0.91 (0.25 to 2.33), based on

four deaths from eight studies. There were no sinonasal cancer deaths and the meta-RR for lung cancer was 0.77 (0.71 to 0.83), based on 562 deaths from eight studies, a statistically significant deficit. From non-nested case-control studies, the meta-RR for nasopharyngeal cancer was 1.3 (0.9 to 2.1), based on 445 exposed cases from six studies. The meta-RR for sinonasal cancer was 1.8 (1.4 to 2.3), based on 933 exposed cases from 11 studies, a statistically significant excess. For lung cancer the meta-RR was 0.8 (0.7 to 1.0), based on 301 exposed cases from six studies, a marginally statistically significant deficit. For nasopharyngeal cancer the authors concluded that the cohort studies did not show an overall excess when corrected for unreported data and that, taken together, the cohort and case-control studies did not provide convincing evidence of a causal relationship between formaldehyde and nasopharyngeal cancer. For sinonasal cancer, they concluded overall that the data did not suggest a relationship with formaldehyde exposure, although many studies did not distinguish between cancers of the nasal cavity and nasal sinuses, since formaldehyde does not seem to penetrate into the nasal sinuses where most human sinonasal cancer originate. The authors concluded that the results of the meta-analysis suggested there was no relationship between lung cancer and formaldehyde exposure [76].

A subsequent sixth meta-analysis examined the evidence for carcinogenicity in cohort studies of industrial and professional workers [121]. The relative risk for nasopharyngeal cancer in industrial cohorts was 1.33 (0.61 to 2.53) based on nine deaths from three studies. For sinonasal cancer, the meta-RR from industrial cohorts was 1.01 (0.33 to 2.35) based on five deaths from three studies and for lung cancer the meta-RR was 1.07 (1.01 to 1.12) based on 1459 cases from six studies, an excess that was statistically significant. Nasopharyngeal cancer was not examined. There was no sinonasal cancer deaths in the studies looked at and the meta-RR for lung cancer was 0.68 (0.63 to 0.74) based on 703 deaths from 13 studies, a statistically significant

deficit. The authors concluded that a comprehensive review of cancer in industry workers and professionals exposed to formaldehyde showed no appreciable excess risk for pharyngeal, sinonasal or lung cancers.

In a more recent seventh meta-analysis of nasopharyngeal cancer, the meta-RR was 0.72 (0.40 to 1.29) based on 39 deaths from 11 cohort studies [125]. From case-control studies, the meta-RR was 1.22 (1.00 to 1.50) based on 502 exposed cases from seven studies. The authors concluded that restricting attention to stronger cohort and case-control study designs, considering the effects of smoking, and ignoring the results from an anomalous plant in the NCI cohort study, their meta-analysis provided little support for a causal relationship between formaldehyde exposure and nasopharyngeal cancer.

The most recent eighth meta-analysis examining the relationship between formaldehyde exposure and sinonasal cancer produced a meta-RR of 1.68 (1.37 to 2.06) for case control and 1.09 (0.66 to 1.79) for cohort studies. The authors concluded that their summary risk estimate strongly suggested that exposure to formaldehyde increased the risk of developing sinonasal cancer.

In terms of the choice of effect estimates to extract from the individual studies, it is interesting to note that Partanen in his meta-analysis [124] criticized the earlier meta-analysis by Blair *et al* [119] for not adopting the general principle of contrasting the risk connected with formaldehyde exposure that exceeded background exposure to that associated with background exposure (so called unexposed). Partanen stated that the predetection latency period was accounted for when lagged inputs (referring to a period of 10 years after the onset of exposure) could be used in the meta-analysis and in addition confounder-adjusted values whenever available.

5.1.8 Motivation and scope of this meta-analysis

The controversy over the so-called anomalous factory in the NCI cohort (see opening section of this Chapter), together with inconsistencies in the findings of the meta-analyses, caused sometimes by different inclusion criteria for the meta-analysis that have been conducted, and the exclusion from quantitative consideration of the cohort studies that have no cases, all emphasise the complex nature of the epidemiological literature. The aim of this meta-analysis was to determine:

- (i) whether there is evidence that occupational exposure to formaldehyde might be associated with an increased risk of lung cancer;
- (ii) how robust is the conclusion in relation to nasopharyngeal cancer determined by IARC [88]; and
- (iii) how robust is the conclusion in relation to sinonasal cancer from the most recent meta-analysis [123].

An updated meta-analysis is also timely, because of recent updates to the three large industrial cohorts [89, 126, 127].

5.2 METHODS

The reporting for this meta-analysis carried out in this Chapter will follow the MOOSE guidelines [128]. The checklist is set out in Table 5.3.

Table 5.3 – MOOSE checklist for meta-analyses in observational epidemiology

Study Characteristic	Items to be reported
Background	<p>Problem definition</p> <p>Hypothesis statement</p> <p>Description of study outcome(s)</p> <p>Type of exposure</p> <p>Type of study designs used</p> <p>Study population</p>
Search strategy	<p>Qualifications of searchers (e.g. librarians, investigators)</p> <p>Search strategy, including time period included in the synthesis and keywords</p> <p>Efforts to include all available studies, including contact with authors</p> <p>Databases and registries searched</p> <p>Search software used, name and version, including special features (e.g. explosion)</p> <p>Use of hand searching (e.g. reference lists of obtained articles)</p> <p>List of citations located and those excluded, including justification</p> <p>Method of addressing articles published in languages other than English</p> <p>Method of handling abstracts and unpublished studies</p> <p>Description of any contact with authors</p>
Methods	<p>Description of relevance or appropriateness of studies assembled for assessing the hypothesis to be tested</p> <p>Rationale for the selection and coding of data</p> <p>Documentation of how data were classified and coded</p> <p>Assessment of confounding</p> <p>Assessment of study quality, including blinding of study assessors; stratification or regression on possible predictors of study results</p> <p>Assessment of heterogeneity</p> <p>Description of statistical methods (e.g. complete description of fixed or random effects models, justification of whether the chosen models account for predictors of study results, dose-response models, or cumulative meta-analysis) in sufficient detail to be replicated.</p>

Study Characteristic	Items to be reported
	Provision of appropriate tables and graphs
Results	<p>Graphic summarizing individual study estimates and overall estimate</p> <p>Table giving description of for each study included</p> <p>Results of sensitivity testing (e.g. subgroup analysis)</p> <p>Indication of statistical uncertainty of findings</p>
Discussion	<p>Quantitative assessment of bias (e.g. publication bias)</p> <p>Justification of exclusion (e.g. exclusion of non-English-language citations)</p> <p>Assessment of quality of included studies (in this case the Newcastle-Ottawa scale [129] for use with cohort and case-control studies)</p>
Conclusions	<p>Consideration of alternative explanations for observed results</p> <p>Generalisation of conclusions (i.e. appropriate for the data presented and within the domain of the literature review)</p> <p>Guidelines for future research</p> <p>Disclosure of funding source</p>

5.2.1 Study identification

Epidemiological studies (no restriction on study design) that examined the risk of nasopharyngeal cancer, sinonasal cancer, or lung cancer (mortality or cancer incidence) from occupational exposure to formaldehyde were originally identified by searching online the databases Medline and Embase covering papers from 1980 to the latest available in August 2002. This included papers published in languages other than English. The searches were carried out in consultation with qualified librarians from the Health and Safety Executive's Information Centre in Sheffield, who had many years of experience of searching scientific databases. Search terms that were used for formaldehyde were: formaldehyde, formalin, fannofom, formalith, formic aldehyde, formol, fyde, lysoform, methaldehyde, methanal, methyl aldehyde, methylene oxide, methyl oxide, morbicid, oxomethane, oxymethylene, paraform or supersyloform.

Occupational studies were identified using the terms: occupation*, work*, employ*, job*, industry*, labor* and labour*. Cancers of the respiratory system were identified using: nasopharynx*, nose, nasal, pharynx*, respira*, thorax*, thorac*, trachea*, chest*, windpipe*, lung*, pulmon*, pneumo*, larynx*, glott*, epiglott*, bronch*, alveoli*, pleura*, and pleuri* combined with cancer*, tumour*, tumor*, neoplasm*, and carcinoma*. Thus potentially relevant papers were identified if they contained at least one of the terms for formaldehyde, occupational studies and a relevant tumour. Citations from the relevant previously published meta-analyses [76, 117-119, 124] were also used, together with any other relevant papers already held. Some papers included in the previous meta-analyses were obtained by contacting the principal author of one of the meta-analyses [76]. Citations were examined for other relevant studies, particularly the bibliographies of two more recent meta-analyses involving formaldehyde and pancreatic cancer and leukaemia [130, 131]. Material not published in the scientific literature and abstracts were not specifically searched for, but were obtained where possible, if cited in a previous meta-analysis.

The literature search was supplemented by a search of PubMed and Medline from 2002 to April 2016 carried out by the author of this work, with advice from a university librarian, using the same search strings as set out above, and inclusion of any additional studies cited in the more recent meta-analyses [121, 123, 125].

5.2.2 Selection criteria for study inclusion in the meta-analysis

Studies were included in the meta-analysis if they contained a risk estimate that allowed inclusion in a meta-analysis of occupational exposure to formaldehyde in relation to at least one of the cancers of interest: nasopharyngeal cancer, sinonasal cancer and lung cancer.

Studies which presented a relative risk without a measure of uncertainty and a means of calculating one were excluded for example Petersen, 1980, Hernberg *et al*, 1983, Zhang & Jiang, 1988 and Krieger 1983 [132-134]. Studies where formaldehyde was not of interest or was not of interest in relation to respiratory cancer such as Dubrow & Wegman, 1984 and Band *et al*, 1997 [135, 136], and studies where the exposure was not occupational for example Friedman & Ury 1983, Vaughan *et al*, 1986 and Walrath *et al*, 1985 [137-139] were also excluded. Studies not examining a cancer site of interest were also excluded such as Zhang & Jiang, 1988[134].

Studies where there was potential for substantial co-exposure to another occupational respiratory carcinogen, but where the risk associated with formaldehyde exposure was not adjusted for these co-exposures were not included, for example Coggon *et al*, 1984, Costanini *et al*, 1989, and Chiazze *et al*, 1997 [140-142].

Studies that were updated by subsequent studies or studies where more than half the study subjects were included in a subsequent study were also excluded for example Marsh *et al* 1996, Fayerweather *et al*, 1983 and Liebling *et al*, 1984 [143-145].

5.2.3 Extraction and coding of study characteristics

Where there was a choice of more than one risk estimate to extract, the following preferences were made:

- Where a subpopulation of a study were assessed as having been the only part of the cohort occupationally exposed to formaldehyde, then results for the exposed sub-cohort were extracted, for example Beane Freeman *et al* [89].
- Where there existed co-exposure to another occupational respiratory carcinogen, but rather than adjust for the co-exposure, the data were

presented for those exposed to formaldehyde in the absence of the co-exposure, then these results were extracted, for example Partanen *et al*, 1990 [146].

- Where lagged and unlagged risk estimates were available, the lagged results, the optimum lag as determined by the authors of the paper, were extracted, for example Partanen *et al*, 1990 [146].
- In the absence of specific results for lung cancer, results for all respiratory cancer combined or for lung and pleura combined was assumed sufficiently relevant to be regarded as lung cancer results for example, Walrath and Fraumeni *et al*, 1984 [147].
- For nested case-control studies, where presented, the results for the nested case-control study were preferred to those of the cohort study, for example Bond *et al*, 1986 [148]
- Where results for PCMRs were presented alongside PMRs, the results based on PCMRs were preferred, although in practice, the percentage difference in RRs was very small, for example, Walrath and Fraumeni, 1983 [149]. (If PMRs are the proportion of deaths due to a specific cause expressed as a proportion of all causes of death, then the PCMR is the same entity, but expressed as a proportion of all deaths due to cancer).
- Confounder-adjusted results were preferred to those unadjusted for confounders, for example, Armstrong *et al*, 2000 [98].

In the absence of an expected number of nasopharyngeal cancer deaths or an expected number of sinonasal cancer deaths in a study, the expected numbers were estimated using the expected number of lung cancer deaths (ELD), by taking their ratio as to the expected number of lung cancer deaths as applicable to occupational cohort studies. For

nasopharyngeal cancer the expected number was taken as $0.004 \times \text{ELD}$ for example in Meyers *et al*, 2013 and for sinonasal cancer the expected number was taken as $0.003 \times \text{ELD}$, for example, Meyers *et al*, 2013 [127]. .

The standard error for the log of the relative risk was estimated as follows:

Study Type	SE(logRR)
Cohort, PMR, Case-control and other studies	$\frac{\log(U95\%CL) - \log(L95\%CL)}{3.92}$
When O = 0.	Undefined

The standard error for the relative risk was estimated as follows:

Study Type	SE(RR)
Cohort studies or PMR studies	$1/\sqrt{E}$
Case-control studies	$RR * SE(\log RR)$ (using the Delta method [78]).

The delta method is an intuitive technique for approximating the moments of functions of random variables and the delta method utilizes a truncated Taylor series expansion [78].

Missing 95% confidence intervals were calculated using an exact method via the Poisson distribution [74], for example in Meyers *et al*, 2013 [127].

5.2.4 Meta-analytical methods

Statistical analyses were performed with the “metan” and “metabias” commands in Stata version 13 [53]. Data utilized when analyses were carried out on the log scale were the natural log-transformed risk ratios and their associated standard errors. For analyses on the original untransformed scale, the data utilized were the risk ratios and their associated standard errors. The coefficients of inconsistency (I^2) were estimated to

assess heterogeneity between studies [150]. I^2 is an estimate of the percentage of total variation in study estimates due to heterogeneity rather than chance, and is considered substantial if it exceeds 50%. When significant heterogeneity is present, the random effects analysis is preferred to the fixed effect analysis. Random effects models [18] were applied to calculate the (reverse transformed) pooled risk ratios and associated 95% CIs and z scores of all studies.

Differences between subgroups, due to study characteristics or study quality, were assessed via subgroup analyses in 'metan'. Study quality was assessed using the Newcastle-Ottawa scale for cohort and case-control studies [129]. Note that weighting studies based on quality scores was not carried out as this can be problematic and produce biased pooled effect estimates [34]. The subgroups defined *a priori* were newer versus older studies, study location, whether risk estimate adjusted for smoking, whether outcome was mortality or incidence, and study quality.

Publication bias was assessed by constructing funnel plots of the log risk ratio versus the standard error of the log risk ratio and of the risk ratio versus the standard error of the risk ratio [151]. To supplement the funnel plot approach, the adjusted rank correlation method suggested by Begg and Mazumdar [27] and the regression asymmetry test proposed by Egger *et al* [26] were applied.

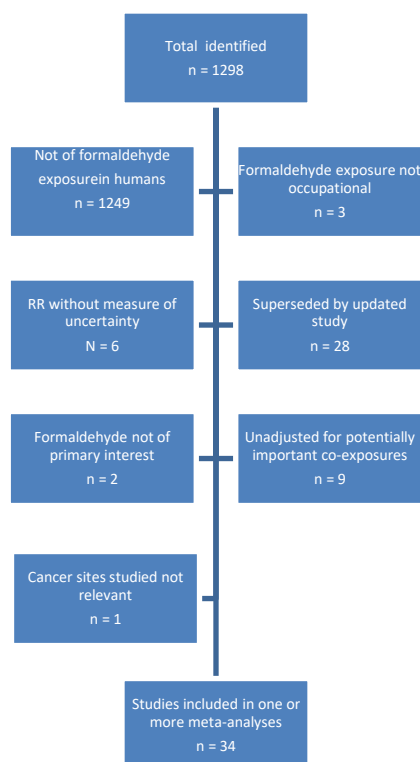
5.3 RESULTS

5.3.1 Characteristics of the studies analysed

Figure 5.1 contains the results of the literature searching. A total of 34 studies giving rise to 73 relative risk estimates of occupational exposure to formaldehyde for a cancer site of interest were available for inclusion in the meta-analyses. Six studies were industrial cohort studies, four were studies of professionals such as pathologists,

anatomists, embalmers and funeral directors, 16 were case-control studies (including one case-cohort study) and eight other study (mainly registry-based) types. The process of identification of the relevant studies is shown diagrammatically in Figure 5.1.

Figure 5.1 - Identification of studies for including the meta-analysis



The characteristics of these studies are set out in Table 5.4 (industrial cohort studies), Table 5.5 (cohort studies of pathologists, anatomists and funeral directors and embalmers, Table 5.6 (nasopharyngeal cancer case control studies), Table 5.7 (sinonasal cancer case-control studies), Table 5.8 (lung cancer case-control studies), and Table 5.9 (other study types).

Table 5.4 – Occupational cohorts of industrial workers

Reference	Location	Number of cases	Study Population	Industry	Follow-up	Source of controls	Variables included in adjustment	Exposure assessment	Newcastle-Ottawa scale score	Effect estimate (95% CI)
Edling <i>et al</i> , 1987 [152]	Sweden	L: 2 N: 1	521	Abrasive manufacture	1958-1983	National population	A, S, Y	Not defined.	6	L: 0.57 (0.1 to 2.1) N: 95.2 (0.0 to 351.3#)
Bertazzi <i>et al</i> , 1989 [153]	Italy	L: 6	1330	Resin manufacture	1959-1986	Local population (not defined)	A, S, Y	Ever exposed, duration	5	L: 0.69 (0.25 to 1.50)
Andjelkovich <i>et al</i> , 1995 [154]	USA	L: 51 N: 0 SN: 0	3,929	Automotive iron foundry workers	1960-1989	National population	A, S, Y	Ever	7	L: 1.20 (0.89 to 1.58) N: 0.00 (0.00 to 23.55) SN: 0.00 (0.00 to 17.62)
Beane Freeman <i>et al</i> , 2013 [89]	USA	N: 9 SN: 3 L: 1,130	25,619	Industrial workers	1934-2004	National population	A, S, Y	Average intensity, peak, cumulative, duration	6	N: 1.84 (0.84 to 3.49) SN: 0.90 (0.18 to 2.62) L: 1.14 (1.07 to 1.20)
Meyers <i>et al</i> , 2013 [127]	USA	N: 0 SN: 0 L: 267	11,043	Garment industry workers	1960-2008	National population	A, S, Y	Ever vs never, duration	6	N: 0.00 (0.00 to 2.77) SN: 0.00 (0.00 to 3.89) L: 1.04 (0.92 to 1.17)
Coggon <i>et al</i> , 2014 [126]	UK	N: 0 SN: 1 L: 813	14,008	Industrial workers	1941-2012	National population	A, S, Y	Highest level, duration	6	N: 0.00 (0.00 to 1.90) SN: 0.52 (0.01 to 2.88) L: 1.26 (1.17 to 1.35)
Abbreviations: A = age; S = sex; Y = year; L = lung cancer; N = nasopharyngeal cancer; SN = sinonasal cancer										
Notes: # Expected value derived using ratio of nasopharyngeal cancers to lung cancers (based on data for England and Wales) applied to the lung cancer expectation										

Table 5.5 – Occupational cohort studies of pathologists, anatomists, funeral directors and embalmers

Reference	Location	Number of cases	Study Population	Industry	Follow-up	Source of controls	Variables included in adjustment	Exposure assessment	Newcastle-Ottawa scale score	Effect estimate (95% CI)
Harrington & Shannon, 1975 [155]	United Kingdom	L: 11	2,079	Pathology	1955-1973	National populations	A, S, Y	None	7	L: 0.39 (0.20 to 0.71)
Levine <i>et al</i> , 1983	Canada	SN: 0 L: 19	1,477 men	Undertakers	1950-1977	National population	A, Y	None	6	SN: 0.0 (0.0 to 15.0) L: 0.94 (0.6 to 1.5)
Stroup <i>et al</i> , 1986 [156]	USA	L: 12 SN: 0 N:	2,317	Anatomy	1925-1979	National population	A, S, Y	None	5	L: 0.3 (0.1 to 0.5) SN: 0.0 (0.0 to 2.0) N: 0.0 (0.0 to 23.2)
Hall <i>et al</i> , 1991 [157]	England and Wales Scotland	L: 9 L: 5	4,512	Pathology	1974-1987	National populations	A, S, Y	None	6	L: 0.19 (0.09 to 0.36) – E&W (England and Wales) L: 0.60 (0.16 to 1.54) – S (Scotland)
Abbreviations: A = age; S = sex; Y = year; L = lung cancer; N = nasopharyngeal cancer; SN = sinonasal cancer										

Table 5.6 – Nasopharygeal cancer case-control studies

Reference	Location	Number of cases	Study Population	Industry	Follow-up	Source of controls	Variables included in adjustment	Exposure assessment	Newcastle-Ottawa scale score	Effect estimate (95% CI)
Roush <i>et al</i> , 1987 [158]	USA	173	976	Population study	1935-1975	Deaths in the region	A, Y, AO	Probability of exposure, level of exposure.	5	1.3 (0.7 to 2.4)
West <i>et al</i> , 1993 [159]	Philippines	104	309	Population-based	Two-year period in late 1980s/early 1990s	Hospital and community	EL, DE, D, TS, AMC, HM	Likely/unlikely exposed	5	1.2 (0.31 to 3.6) =- <15y (lag less than 15 years) 4.0 (1.3 to 12.3) – 15+ (lag 15 years or more)
Armstrong <i>et al</i> , 2000 [98]	Malaysia	530	1060	Population-based	1990-1992	Community	D, TS	Ever exposed	6	0.71 (0.34 to 1.43)
Vaughan <i>et al</i> , 2000 [160]	USA	196	440	Population-based	1987-1993	Community	A, S, E, LN, TS, AL, EL	Probability, maximum, duration, cumulative	6	1.3 (0.8 to 2.1)
Hildesheim <i>et al</i> , 2001	Taiwan	375	702	Population-based	1991-1994	Community	A, S, EL, E, HLA, FH, TS, G, D = diet	Probability, intensity, duration	6	1.4 (0.93 to 2.2)
Abbreviations: A = age; E = ethnicity; Y = year; AO = availability of occupational data; S = Sex; EL = Education level; DE = dust/exhaust; D = diet; TS = tobacco smoking, AMC = anti-mosquito coils, HM = herbal medicine, E = ethnicity, HLA = HLA allele, FH = family history, G = genotype, D = diet, LN = location, AL = alcohol										

Table 5.7 – Sinonasal cancer case-control studies

Reference	Location	Number of cases	Study Population	Industry	Follow-up	Source of controls	Variables included in adjustment	Exposure assessment	Newcastle-Ottawa scale score	Effect estimate (95% CI)
Olsen & Anaes, 1986 [161]	Denmark	2 (squamous cell carcinoma) 1 (adenocarcinoma)	2465	Population study	1970-1982	Cancer registry	W	Ever exposed (hygiene assessment based on job histories)	4	1.4 (0.3 to 6.4) – SCC (squamous cell carcinoma) 9.5 (1.6 to 57.8) – A (adenocarcinoma)
Roush <i>et al</i> , 1987 [158]	USA	198	976	Population study	1935-1975	Deaths in the region	A, Y, AO	Probability of exposure, level of exposure.	5	1.0 (0.5 to 1.8)
Luce <i>et al</i> , 2002 [162]	Various (pooled analysis)	1,213	4,349	Various	1970-1990	Hospital, Community	A, S, SY	Level of exposure	7	1.2 (0.8 to 1.8) – M SCC low (Male squamous cell carcinoma low exposure) 1.1 (0.8 to 1.6) – M SCC med (Male squamous cell carcinoma medium) 1.2 (0.8 to 1.8) – M SCC high (Male squamous cell carcinoma high) 0.7 (0.3 to 1.9) – M A low (male adenocarcinoma low) 2.4 (1.3 to 4.5) – M A med (male adenocarcinoma medium) 3.0 (1.5 to 5.7) – M A high (male adenocarcinoma high) 0.6 (0.2 to 1.4) – F SCC low (Female squamous cell carcinoma low) 1.3 (0.6 to 3.2) – F SCC

Reference	Location	Number of cases	Study Population	Industry	Follow-up	Source of controls	Variables included in adjustment	Exposure assessment	Newcastle-Ottawa scale score	Effect estimate (95% CI)
										med (Female squamous cell carcinoma medium) 1.5 (0.6 to 3.8) – F SCC high (Female squamous cell carcinoma high) 2 (0.2 to 4.1) – F A low (Female adenocarcinoma low) 6.2 (2.0 to 19.7) – F A high (Female adenocarcinoma high)
Pesch <i>et al</i> , 2008 [163]	Germany	86	290	Industry-based	2003-2005	Insurance database	A, R, RS, TS	Ever versus never	6	0.46 (0.24 to 1.54) 0.94 (0.47 to 1.90)
Abbreviations: W = wood dust; A = age; Y = year; AO = availability of occupational information; R = region; RS = respondent status; SY = study; TS = tobacco smoking										

Table 5.8 – Lung cancer case-control studies

Reference	Location	Number of cases	Study Population	Industry	Follow-up	Source of controls	Variables included in adjustment	Exposure assessment	Newcastle-Ottawa scale score	Effect estimate (95% CI)
Bond <i>et al</i> , 1986 [148]	USA	4	19,608	Chemical	1940-1981	Decedent and live controls from within cohort	TS, V, YM, DR, EL	Ever exposed	6	0.31 (0.11 to 0.86)
Gerin <i>et al</i> , 1989 [164]	Canada	857	2,380	Population-based	1979-1985	General population	A, E, SES, TS, DJ	Ever exposed, probability of exposure, frequency of exposure, concentration, duration, era of first exposure (hygiene assessment based on job histories)	7	1.0 (0.6 to 1.8) - Short (short exposed) 0.5 (0.3 to 0.8) - Long_low (long low exposure) 0.9 (0.5 to 1.6) Long_med (long medium exposure) 1.0 (0.4 to 1.0) – long_high (long high exposure)
Partanen <i>et al</i> , 1990 [146]	Finland	6	7307	Woodworking	1957-1982	Cohort members without respiratory cancer	VS, TS	Ever exposed, duration, level of exposure, cumulative exposure, repeated peak exposure (hygiene assessment based on job histories)	7	1.19 (0.31 to 4.56)
Brownson <i>et al</i> , 1993 [165]	USA	429	1,450	Population-based	1986-1991	Community	TS, PH	None	7	0.9 (0.2 to 3.3)
De Stefani <i>et</i>	Uruguay	338	1,352	Population-	1994-2000	Hospital	TS	Duration	6	1.7 (1.1 to 2.8)

Reference	Location	Number of cases	Study Population	Industry	Follow-up	Source of controls	Variables included in adjustment	Exposure assessment	Newcastle-Ottawa scale score	Effect estimate (95% CI)
<i>al</i> , 2005 [166]				based						
Checkoway <i>et al</i> , 2011 [167] (case-cohort)	China	628	267,400	Textile workers	1989-1998	Referent subcohort	A, TS, EN	Duration	5	2.1 (0.4 to 11)
Mahboubi <i>et al</i> , 2013 [168]	Canada	2060	4,106	Population-based	1979-1996	Community	A, S, IL, E, RS, EL, TS, LC	Age at first, average, peak	7	1.06 (0.89 to 1.27)
Abbreviations: VS = vital status, S = smoking; E = ethnicity; SES = socio-economic status; DJ = dirty jobs; V = vitamin A; YM = year of migration; DR = duration of residence; EL = education level, EN = endotoxin; IL = income level, RS = respondent status, LC = recognized lung carcinogens; PH = previous history of lung disease										

Table 5.9 – Other study types

Reference	Study Type	Location	Number of cases	Study Population	Industry	Follow-up	Source of controls	Variables included in adjustment	Exposure assessment	Effect estimate (95% CI)
Jensen & Andersen, 1982 [169]	Registry-based	Denmark	L: 8 SN: 0	302	Pathology	1943-76	Cancer registry	A, S, Y	None	L: 1.0 (0.4 to 2.4) SN: 0.0 (0.0 to 93.6)
Milham, 1983[170]	Registry-based	USA	L: 13	454,992	Funeral directors	1950-79	Mortality registry	A, S, Y	None	L: 0.75 (0.40 to 1.28)
Walrath & Fraumeni, 1983 [149]	PMR study	USA	L: 72	1,132	Embalmers	1925-1980	National population	A, S, Y	Approximate length of exposure	L: 1.08 (0.84 to 1.36)
Walrath & Fraumeni, 1984 [147]	SMR study	USA	L: 41 SN: 0	1,109	Embalmers	1925-1980	National population	A, S, Y	Length of licensure	L: 0.96 (0.69 to 1.30) SN: 0.00 (0.00 to 4.99)
Hayes <i>et al</i> , 1990 [171]	PMR study	USA	L: 285 L: 23 SN: 0 SN: 0 N: 3 N: 1	3,866	Funeral directors/embalmers	1975-1985	National population	A, S, Y	None	L: 0.97 (0.86 to 1.09) – W (white) L: 0.75 (0.47 to 1.13) – NW (non-white) SN: 0.00 (0.00 to 2.00) – W (white) SN: 0.0 (0.0 to 18.77) – NW (non-white) N: 1.89 (0.39 to 5.48) – W (white) N: 4.00 (0.10 to 22.29) – NW (non-white)
Hansen & Olsen, 1995 [172]	PMR study (incidence)	Denmark	N: 4 SN: 13 L: 410	126,347	Population-based	1970-1984	National population	A, S, Y, White vs blue collar	Low/High exposed	N: 1.3 (0.3 to 3.2) SN: 2.3 (1.3 to 4.0) L: 1.0 (0.9 to 1.1)

Reference	Study Type	Location	Number of cases	Study Population	Industry	Follow-up	Source of controls	Variables included in adjustment	Exposure assessment	Effect estimate (95% CI)
Stellman <i>et al</i> , 1998 [173]	Population-based cohort	USA	L: 104	1.2 million	Population-based	1982-1989	National population	A, TS	Ever versus never	L: 0.93 (0.73 to 1.18)
Siew <i>et al</i> , 2012 [174]	Registry study	Finland	N: 292 SN: 149 L: 30,137	1.2 million	Population-based	1971-1995	National population	A, TS, Silica, Asbestos	Probability, Level	N: 0.87 (0.34 to 2.20) SN: 1.11 (0.66 to 1.87) L: 1.18 (1.12 to 1.25)
Abbreviations: A = age; S = Sex; Y = year; TS = tobacco smoking; L = lung cancer; N = nasopharyngeal cancer; SN = sinonasal cancer										

5.3.2 Lung Cancer

For lung cancer there were 30 estimates of relative risk incorporating 6,476 lung cancers. The results of the fixed effect and random effects meta-analyses are set out in Table 5.10.

Table 5.10 – Summary risks for meta-analyses of all studies of lung cancer and subgroups – analyses on the log scale

Analysis	No. of risk estimates	Summary fixed effect RR (95% CI)	Z (p value)	Summary random effects RR (95% CI)	Z (p value)	Heterogeneity I ² (p-value)	Test of heterogeneity between subgroups	References
All Risk Estimates	30	1.12 (1.08 to 1.15)	7.41 (<0.001)	0.98 (0.90 to 1.07)	0.52 (0.604)	72.8% (<0.001)		Edling <i>et al</i> , 1987; Bertazzi <i>et al</i> , 1989; Andjelkovich <i>et al</i> , 1995; Beane Freeman <i>et al</i> , 2013; Meyers <i>et al</i> , 2013; Coggon <i>et al</i> , 2014; Harrington & Shannon, 1975; Levine <i>et al</i> , 1993; Stroup <i>et al</i> , 1986; Hall <i>et al</i> , 1991; Bond <i>et al</i> , 1986; Gerin <i>et al</i> , 1989; Partanen <i>et al</i> , 1990; Brownson <i>et al</i> , 1993; De Stefani <i>et al</i> , 2005; Checkoway <i>et al</i> , 2011; Mahboubi <i>et al</i> , 2013; Jensen & Anderson, 1982; Milham, 1983; Walrath & Fraumeni 1983; Walrath & Fraumeni, 1984; Hayes <i>et al</i> , 1990; Hansen & Olsen, 1995; Stellman <i>et al</i> , 1998; Siew <i>et al</i> , 2012.
Published before 1990	14	0.84 (0.73 to 0.96)	2.51 (0.012)	0.74 (0.59 to 0.93)	2.61 (0.009)	52.0% (0.012)	< 0.001	Edling <i>et al</i> , 1987; Bertazzi <i>et al</i> , 1989; Harrington & Shannon, 1975; Levine <i>et al</i> , 1983; Stroup <i>et al</i> , 1986; Bond <i>et al</i> , 1996; Gerin <i>et al</i> , 1989; Jensen & Andersen, 1983; Milham, 1983; Walrath & Fraumeni, 1983; Walrath & Fraumeni, 1984;

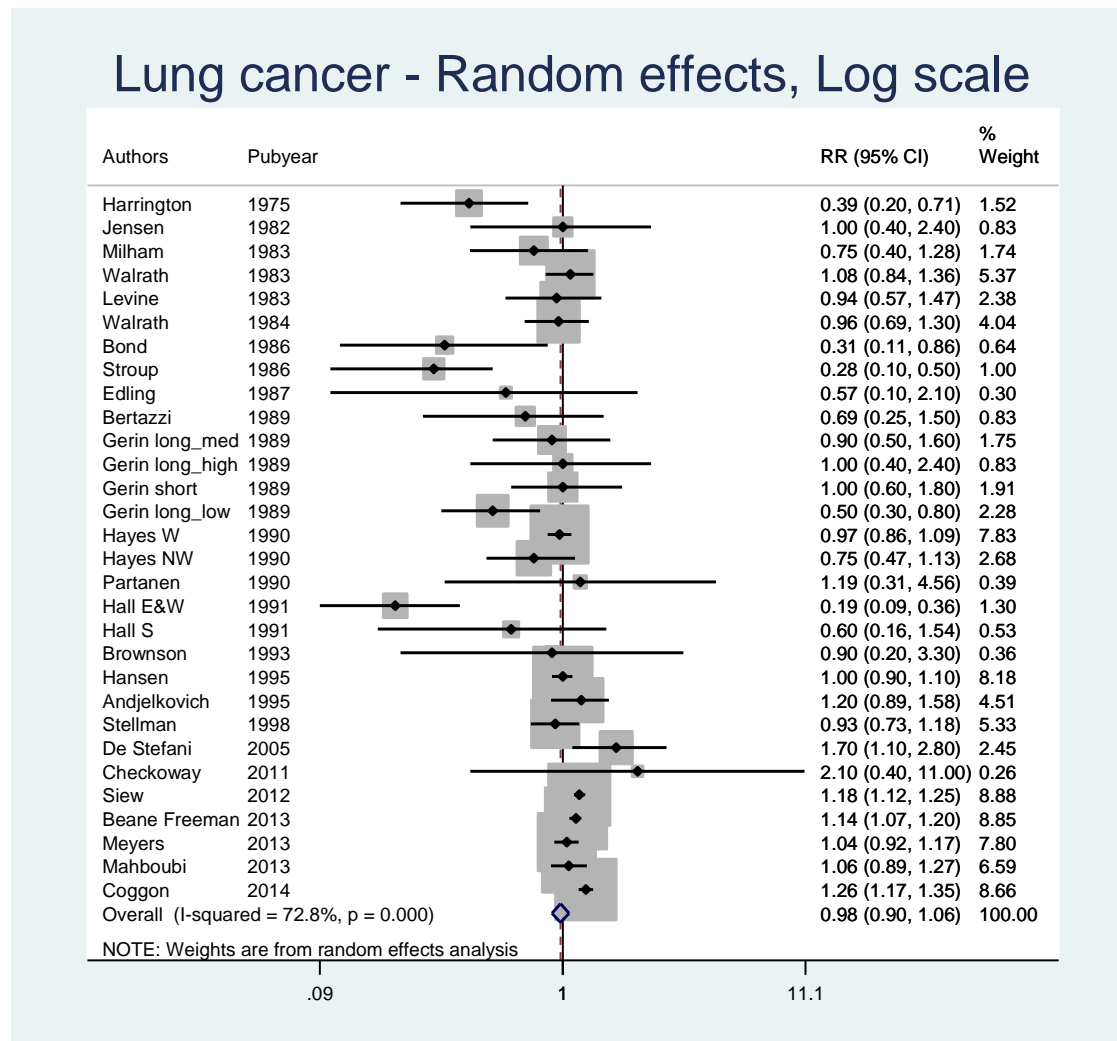
Analysis	No. of risk estimates	Summary fixed effect RR (95% CI)	Z (p value)	Summary random effects RR (95% CI)	Z (p value)	Heterogeneity I ² (p-value)	Test of heterogeneity between subgroups	References
Published 1990 or later	16	1.13 (1.10 to 1.16)	8.13 (<0.001)	1.06 (0.98 to 1.16)	1.43 (0.153)	75.8% (< 0.001)		Andjelkovich <i>et al</i> , 1995; Beane Freeman <i>et al</i> , 2013; Meyers <i>et al</i> , 2013; Coggon <i>et al</i> , 2014; Hall <i>et al</i> , 1991; Partanen <i>et al</i> , 1990; Brownson <i>et al</i> , 1993; De Stefani <i>et al</i> , 2005; Checkoway <i>et al</i> , 2011; Mahboubi <i>et al</i> , 2013; Hayes <i>et al</i> , 1990; Hansen & Olsen, 1995; Stellman <i>et al</i> , 1998; Siew <i>et al</i> , 2012
Study design – industrial cohort	6	1.16 (1.11 to 1.21)	6.90 (< 0.001)	1.04 (0.89 to 1.21)	0.43 (0.669)	81.7% (< 0.001)	< 0.001	Edling <i>et al</i> , 1987; Bartazzi <i>et al</i> , 1989; Andjelkovich <i>et al</i> , 1995; Beane Freeman <i>et al</i> , 2013; Myers <i>et al</i> , 2013; Coggon <i>et al</i> , 2014
Study design – professional cohort	4	0.58 (0.41 to 0.82)	3.08 (0.002)	0.49 (0.23 to 1.09)	1.85 (0.065)	76.7% (0.014)		Harrington & Shannon, 1975; Levine <i>et al</i> , 1983; Stroup <i>et al</i> , 1986; Hall <i>et al</i> , 1991
Study design -case-control	10	1.01 (0.87 to 1.16)	0.08 (0.937)	0.93 (0.70 to 1.25)	0.45 (0.650)	52.6% (0.025)		Bond <i>et al</i> , 1986; Gerin <i>et al</i> , 1989; Partanen <i>et al</i> , 1990; Brownson <i>et al</i> , 1993; De Stefani <i>et al</i> , 2005; Checkoway <i>et al</i> , 2011; Mahboubi <i>et al</i> , 2013
Study design - other	9	1.09 (1.05 to 1.14)	4.16 (< 0.001)	1.01 (0.91 to 1.12)	0.21 (0.835)	62.3% (0.007)		Jensen & Anderson, 1982; Milham, 1983; Walrath & Fraumeni, 1983; Walrath & Fraumeni, 1984; Hayes <i>et al</i> , 1990; Hansen & Olsen, 1995; Stellman <i>et al</i> , 1988; Siew <i>et al</i> , 2012

Analysis	No. of risk estimates	Summary fixed effect RR (95% CI)	Z (p value)	Summary random effects RR (95% CI)	Z (p value)	Heterogeneity I ² (p-value)	Test of heterogeneity between subgroups	References
Outcome - mortality	19	1.11 (1.07 to 1.15)	5.46 (< 0.001)	0.91 (0.81 to 1.03)	1.49 (0.136)	78.2% (< 0.001)	0.654	Edling <i>et al</i> 1987; Bertazzi <i>et al</i> , 1989; Andjelkovich <i>et al</i> , 1995; Beane Freeman <i>et al</i> , 2013; Meyers <i>et al</i> , 2013; Coggon <i>et al</i> , 2014; Harrington & Shannon, 1975; Levine <i>et al</i> , 1983; Stroup <i>et al</i> , 1986; Hall <i>et al</i> , 1991; Bond <i>et al</i> , 1986; Jensen & Andersen, 1982; Milham, 1983; Walrath & Fraumeni, 1983; Walrath & Fraumeni, 1984; Hayes <i>et al</i> , 1990; Stellman <i>et al</i> , 1998.
Outcome - cancer incidence	11	1.12 (1.07 to 1.18)	5.03 (< 0.001)	1.05 (0.92 to 1.20)	0.77 (0.443)	57.6% (0.009)		Gerin <i>et al</i> , 1989; Partanen <i>et al</i> , 1990; Brownson <i>et al</i> , 1993; De Stefani <i>et al</i> , 2005; Checkoway <i>et al</i> , 2011; Mahboubi <i>et al</i> , 2013; Hansen & Olsen, 1994; Siew <i>et al</i> , 2012.
Geographic area - North America	18	1.07 (1.02 to 1.11)	2.92 (0.004)	0.97 (0.88 to 1.06)	0.71 (0.477)	57.7% (0.001)	0.005	Andjelkovich <i>et al</i> , 1995; Bean Freeman <i>et al</i> , 2013; Meyers <i>et al</i> , 2013; Levine <i>et al</i> , 1983; Stroup <i>et al</i> , 1986; Bond <i>et al</i> , 1986; Gerin <i>et al</i> , 1989; Brownson <i>et al</i> , 1993; Mahboubi <i>et al</i> , 2013; Milham, 1983; Walrath & Fraumeni, 1983; Walrath & Fraumeni, 1984; Hayes <i>et al</i> , 1990; Stellman <i>et al</i> , 1998.

Analysis	No. of risk estimates	Summary fixed effect RR (95% CI)	Z (p value)	Summary random effects RR (95% CI)	Z (p value)	Heterogeneity I ² (p-value)	Test of heterogeneity between subgroups	References
Geographic area - Europe	10	1.16 (1.11 to 1.21)	7.24 (< 0.001)	0.93 (0.78 to 1.11)	0.78 (0.434)	83.5% (< 0.001)		Edling <i>et al</i> , 1987; Bertazzi <i>et al</i> , 1989; Coggon <i>et al</i> , 2014; Harrington & Shannon, 1975; Hall <i>et al</i> , 1991; Partanen <i>et al</i> , 1990; Jensen & Andersen, 1982; Hansen & Olsen, 1995; Siew <i>et al</i> , 2012.
Adjusted for smoking - yes	22	1.00 (0.88 to 1.13)	0.02 (0.984)	0.98 (0.79 to 1.21)	0.21 (0.836)	48.5% (0.059)	0.071	Edling <i>et al</i> , 1987; Bertazzi <i>et al</i> , 1989; Andjelkovich <i>et al</i> , 1995; Beane Freeman <i>et al</i> , 2013; Meyers <i>et al</i> , 2013; Coggon <i>et al</i> , 2014; Harrington & Shannon, 1975; Levine <i>et al</i> , 1983; Stroup <i>et al</i> , 1986; Hall <i>et al</i> , 1991. Bond <i>et al</i> , 1986; Brownson <i>et al</i> , 1993; Checkoway <i>et al</i> , 2011; Jensen & Andersen, 1982; Milham 1983; Walrath & Fraumeni, 1983; Walrath & Fraumeni, 1984; Hayes <i>et al</i> , 1990; Hansen & Olsen, 1995; Siew <i>et al</i> , 2012.
Adjusted for smoking - no	8	1.12 (1.09 to 1.16)	7.63 (< 0.001)	0.98 (0.89 to 1.08)	0.43 (0.664)	76.6% (< 0.001)		Mahboubi <i>et al</i> , 2013; Gerin <i>et al</i> , 1989; De Stefani <i>et al</i> , 2005; Stellman <i>et al</i> , 1998; Partanen <i>et al</i> , 1990.
Newcastle-Ottawa scale score - 7	8	0.98 (0.86 to 1.11)	0.35 (0.728)	0.86 (0.66 to 1.13)	1.09 (0.275)	61.1% (0.012)	0.001	Harrington & Shannon, 1975; Gerin <i>et al</i> , 1989; Partanen <i>et al</i> , 1990; Andjelkovich <i>et al</i> , 1995; Mahboubi <i>et al</i> , 2013
Newcastle-Ottawa scale score - 6	10	1.16 (1.11 to 1.21)	6.86 (< 0.001)	1.01 (0.86 to 1.20)	0.17 (0.868)	80.7% (< 0.001)		Levine <i>et al</i> , 1983; Bond <i>et al</i> , 1986; Edling <i>et al</i> , 1987; Hall <i>et al</i> , 1991; Brownson <i>et al</i> , 1993; De Stefani <i>et al</i> , 2005; Meyers <i>et al</i> , 2013; Beane Freeman <i>et al</i> , 2013; Coggon <i>et al</i> , 2014
Newcastle-Ottawa scale score - 5	3	0.50 (0.29 to 0.88)	2.39 (0.017)	0.61 (0.23 to 1.66)	0.97 (0.332)	63.0% (0.067)		Stroup <i>et al</i> , 1986; Bertazzi <i>et al</i> , 1989; Checkoway <i>et al</i> , 2011.

The forest plot for the random effects model is shown in Figure 5.2.

Figure 5.2 – Random effects meta-analysis of lung cancer for all risk estimates

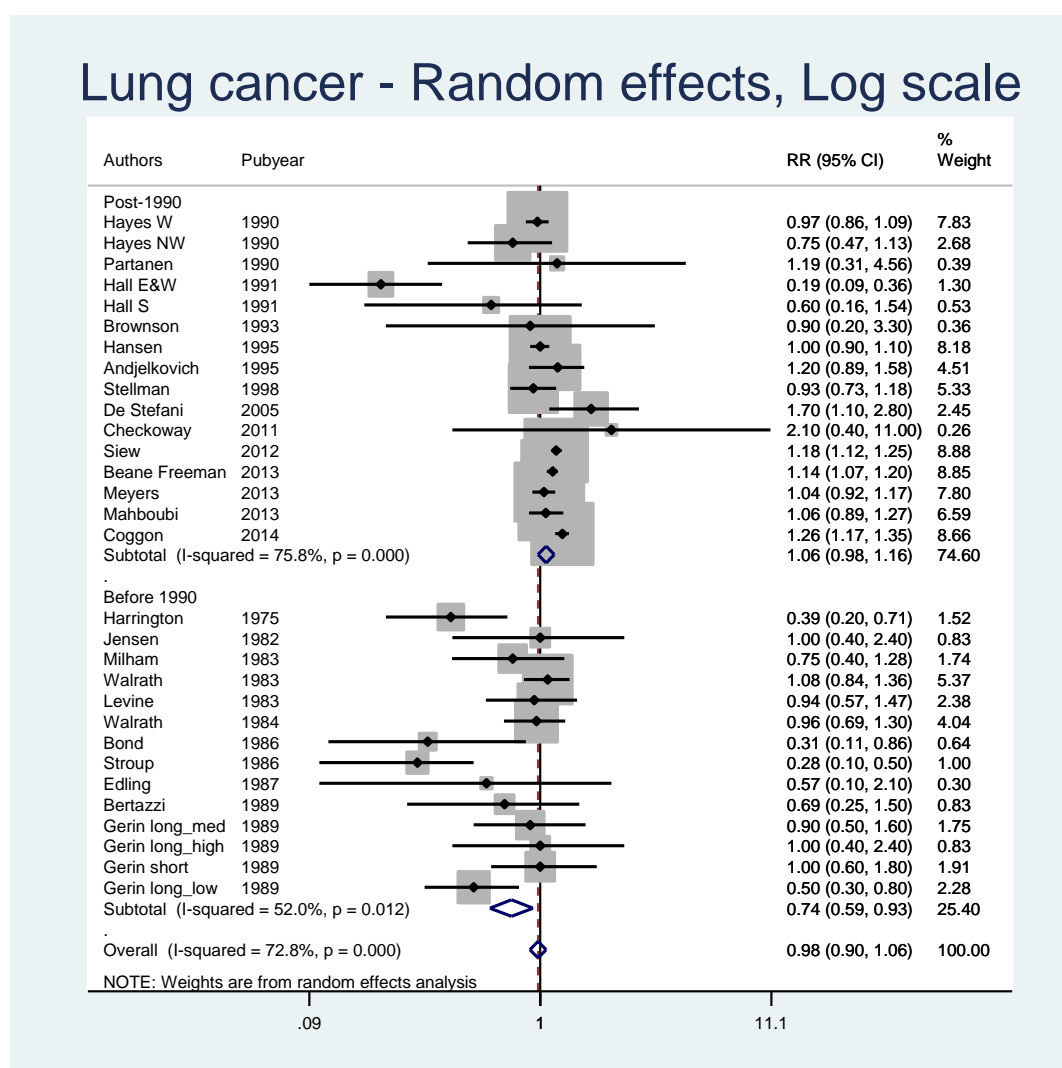


The overall summary or meta-RR for lung cancer is 1.12 (1.08 to 1.15) based on the fixed effect meta-analyses. However, there was significant heterogeneity, and so the random effects meta-analysis is to be preferred. This gave a meta-RR of 0.98 (0.90 to 1.07) which is no longer statistically significantly elevated.

Table 5.10 also contains analysis of important subgroups of studies. The first of these shows that there is a difference in the meta-RR based on the studies published before 1990 and that based on those published in 1990 or later. The meta-RR for the random effects model for pre-1990 publications which gave a meta-RR of 0.74 (0.59 to 0.93),

which was statistically significantly low, whereas the random effect meta-RR for studies published in 1990 or later was 1.06 (0.98 to 1.16) and so was not statistically significantly different to one. The forest plot illustrating this is shown in Figure 5.3.

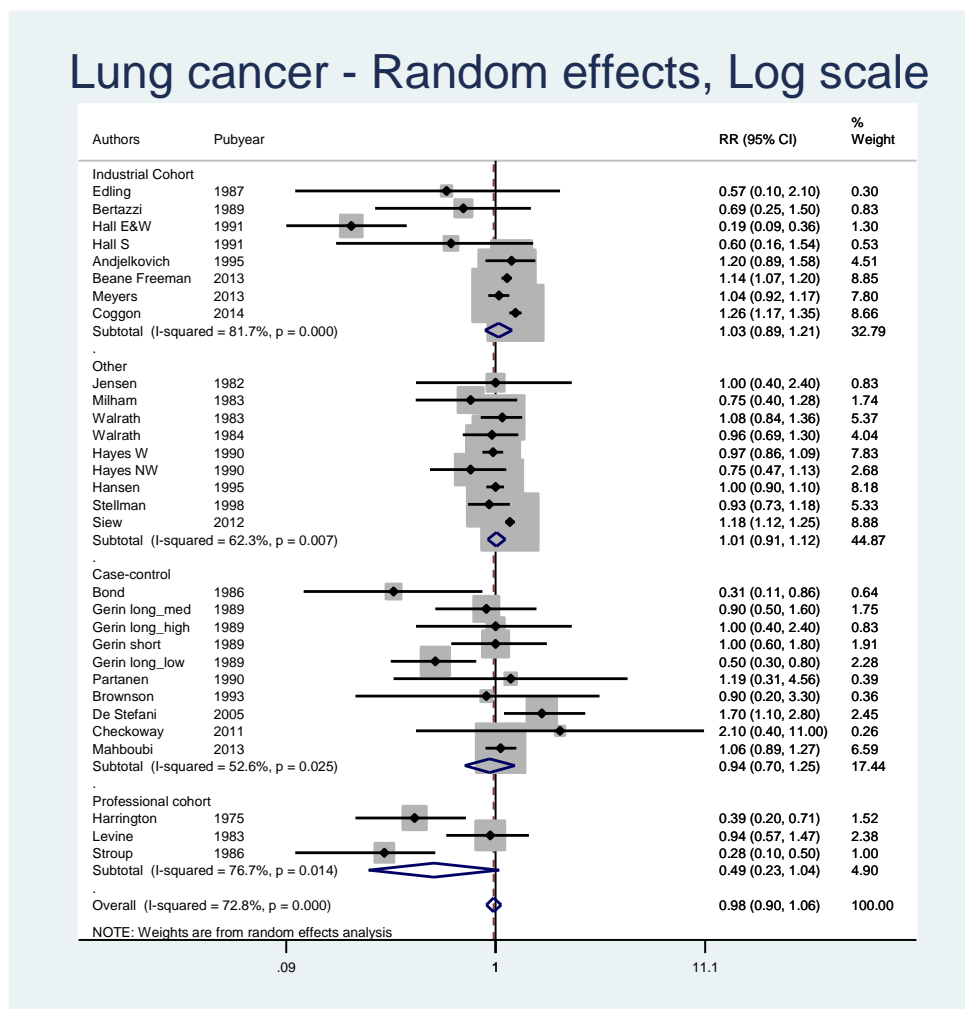
Figure 5.3 – Random effects meta-analysis of lung cancer by year of publication category



Examining the results by study design gave interesting and contrasting meta-RR estimates. Again from the random effect model, the meta-RR for professional cohort studies was 0.49 (0.23 to 1.01) which was borderline statistically significantly low, whereas that based on industrial cohorts was 1.04 (0.89 to 1.21) which not statistically significantly different from one. Case-control studies had a meta-RR of 0.93 (0.70 to 1.25) which was also not statistically significantly different from one. Other studies, such as PMR studies, which are thought to be more prone to bias, gave a meta-RR of

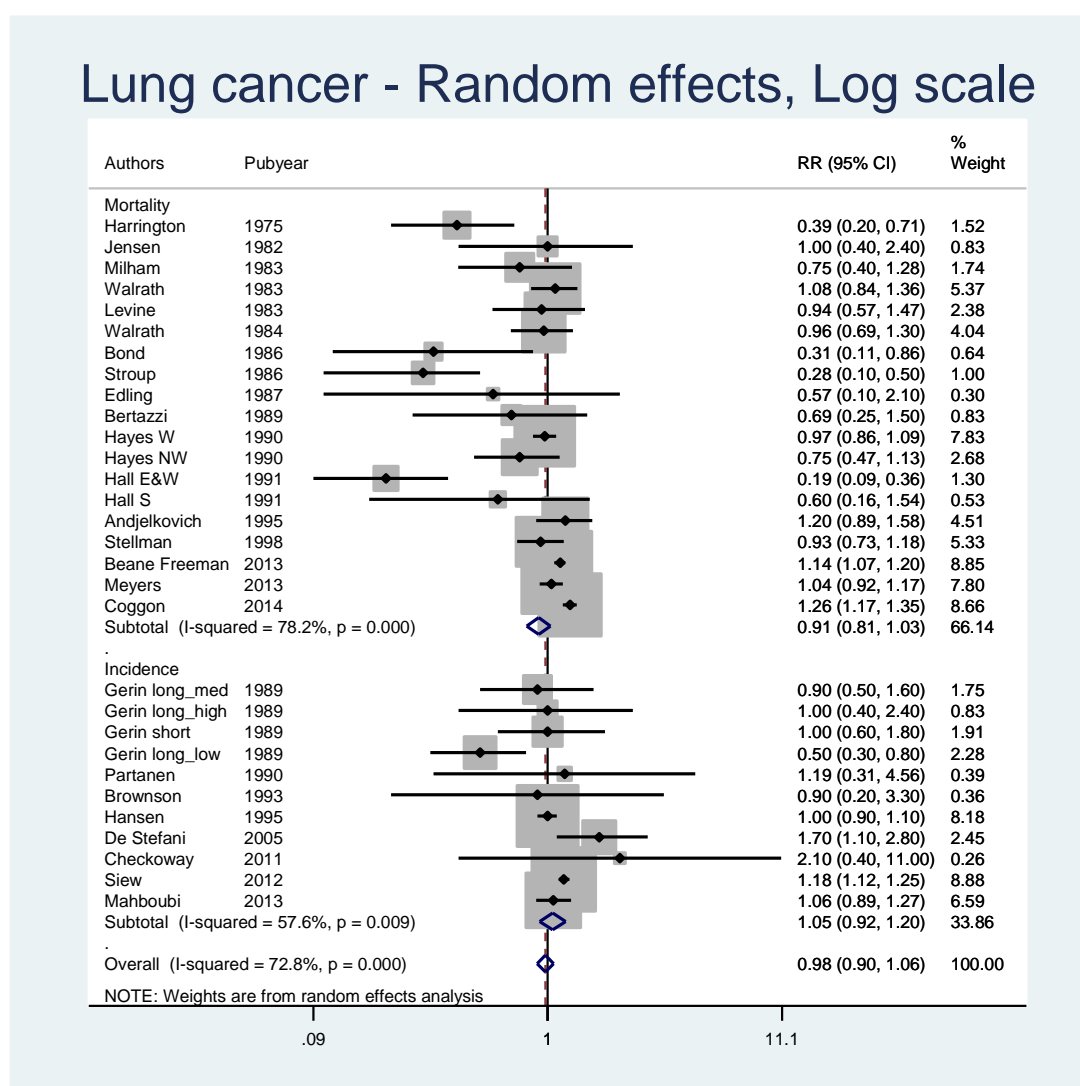
1.01 (0.91 to 1.12) which was also not statistically significantly different to one. This is illustrated in Figure 5.4.

Figure 5.4 - Random effects meta-analysis of lung cancer by study type



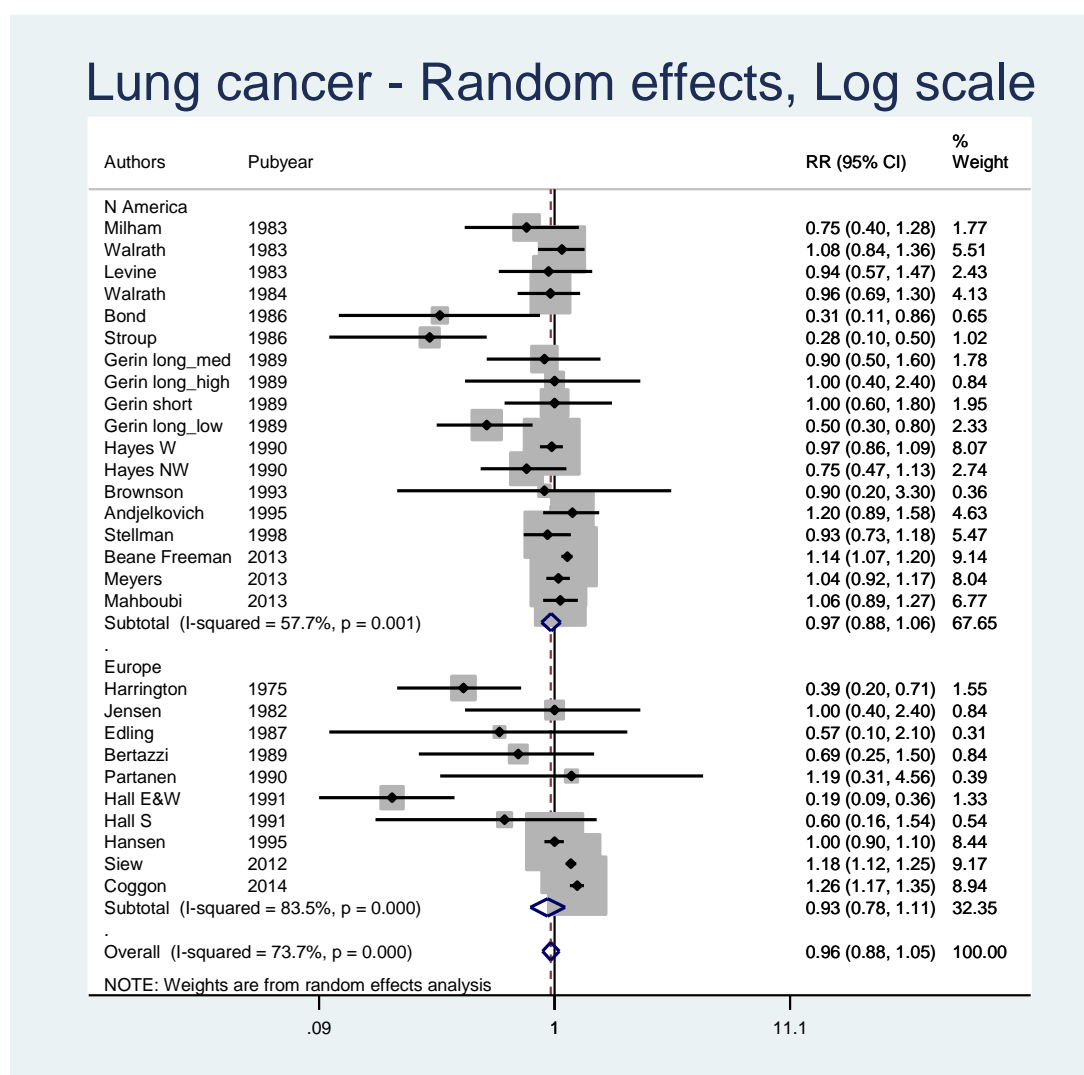
As anticipated for a fairly fatal cancer such as lung cancer, from the random effects model, the meta-RR from mortality studies 0.91 (0.81 to 1.03) gave a similar meta-RR estimate to studies based on cancer incidence data 1.05 (0.92 to 1.20). The forest plot for this analysis is shown in Figure 5.5.

Figure 5.5 - Random effects meta-analysis of lung cancer by mortality or incidence study



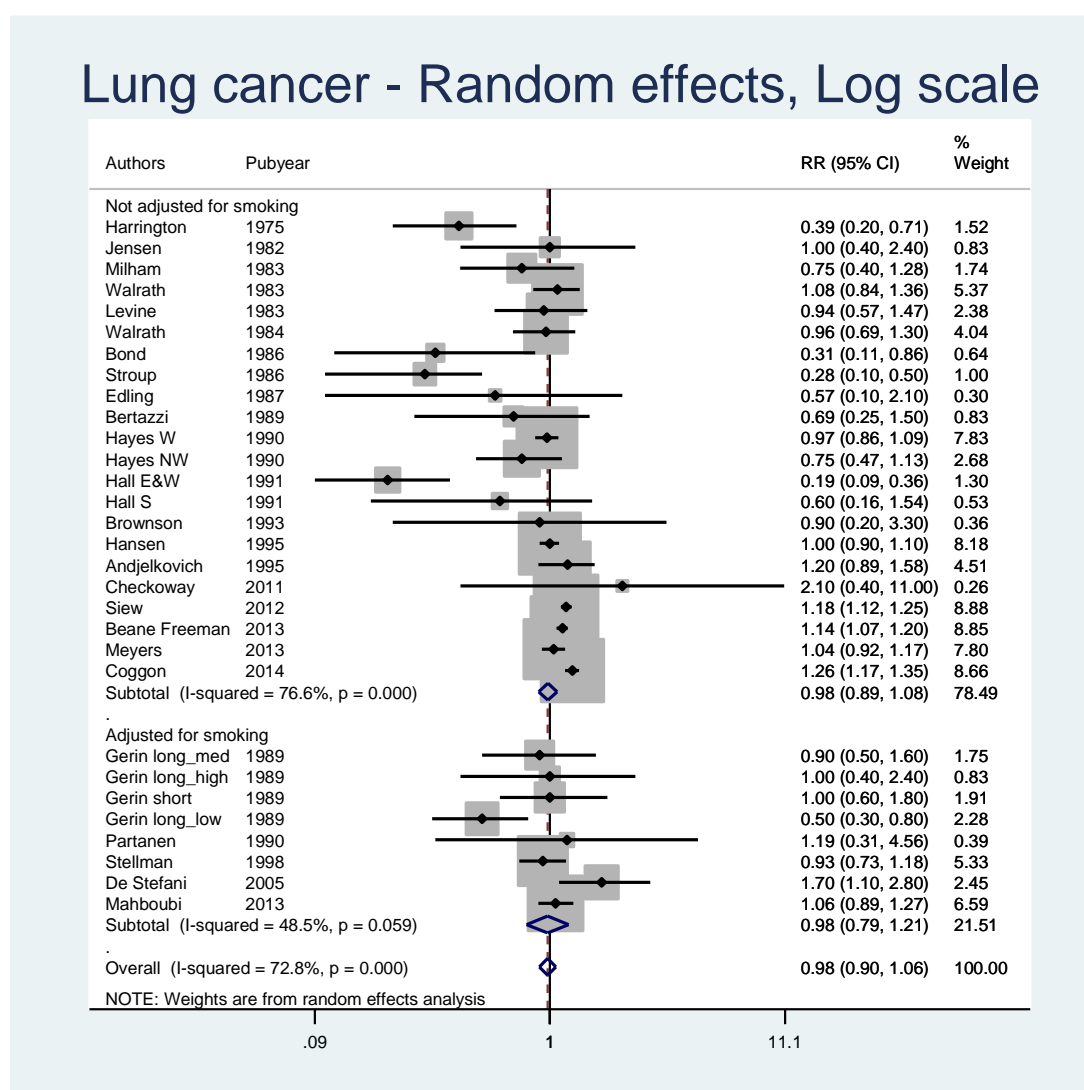
Studies from North America, again using the random effects model, had a meta-RR of 0.97 (0.88 to 1.06) and for Europe the meta-RR was 0.93 (0.78 to 1.11). (One study from South America and one containing studies from multiple countries were excluded from this comparison). The forest plot for this analysis is contained in Figure 5.6.

Figure 5.6 - Random effects meta-analysis of lung cancer by study location



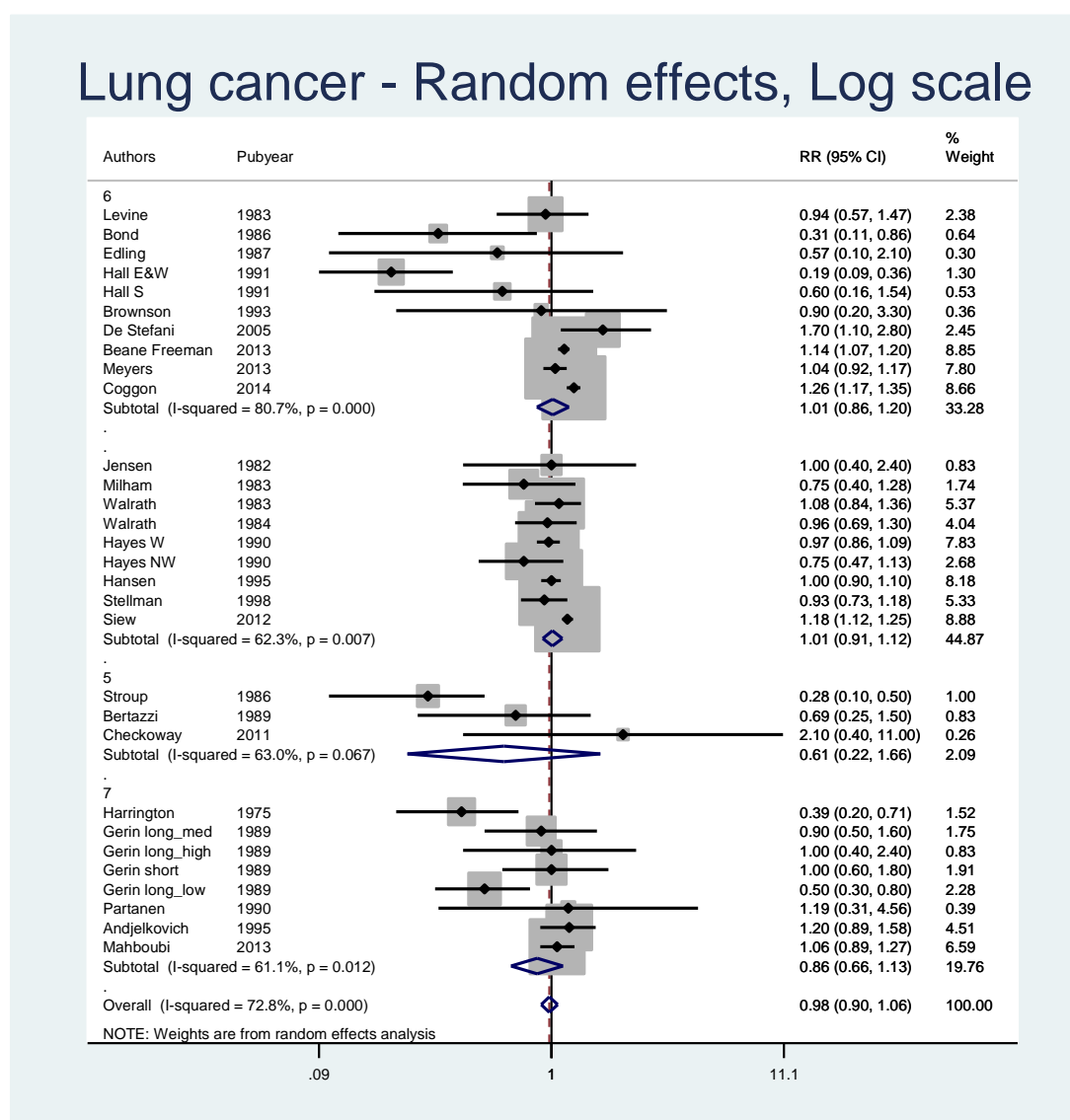
Studies that contained an adjustment for smoking gave a meta-RR of 0.98 (0.79 to 1.21) and studies that did not adjust for smoking a meta-RR 0.98 (0.89 to 1.08), provided some evidence that smoking was not a strong confounder for lung cancer in these studies. The forest plot for this analysis is contained in Figure 5.7.

Figure 5.7 - Random effects meta-analysis of lung cancer by whether or not the study adjusted for smoking



The final subgroup analysis was based on quality score, according to the Newcastle-Ottawa scale. This was only applied to the occupational cohort studies and the case-control studies. The studies with a Newcastle-Ottawa scale score of 7, had a meta-RR of 0.86 (0.66 to 1.13). The meta-RR for studies with a Newcastle-Ottawa scale score of 6 was 1.01 (0.86 to 1.20). The meta-RR for studies with a score of 5 was 0.61 (0.23.to 1.66) albeit based on only three estimates. The forest plot for this analysis is contained in Figure 5.8.

Figure 5.8 – Random effects meta-analysis of lung cancer by study quality category



The lung cancer data displayed considerable heterogeneity, with the value of I^2 being greater than 50% for the vast majority of the subgroups examined.

A sensitivity analysis excluding each study in turn from the random effects analysis is set out in Table 5.11.

Table 5.11 – Meta-RR for random effects model excluding each study in turn

Study Excluded	Meta-RR	95% CI
	0.98	0.90 to 1.07
Edling	0.98	0.90 to 1.07
Bertazzi	0.98	0.90 to 1.07
Andjelkovich	0.97	0.89 to 1.06
Beane Freeman	0.95	0.86 to 1.04
Meyers	0.97	0.88 to 1.06
Coggon	0.95	0.87 to 1.04
Harrington	1.00	0.92 to 1.08
Levine	0.98	0.90 to 1.07
Stroup	1.00	0.92 to 1.08
Hall	1.02	0.94 to 1.10
Bond	0.99	0.91 to 1.07
Gerin	1.00	0.92 to 1.09
Partanen	0.98	0.90 to 1.06
Brownson	0.98	0.90 to 1.07
De Stefani	0.97	0.89 to 1.05
Checkoway	0.98	0.98 to 1.06
Mahboubi	0.97	0.89 to 1.06

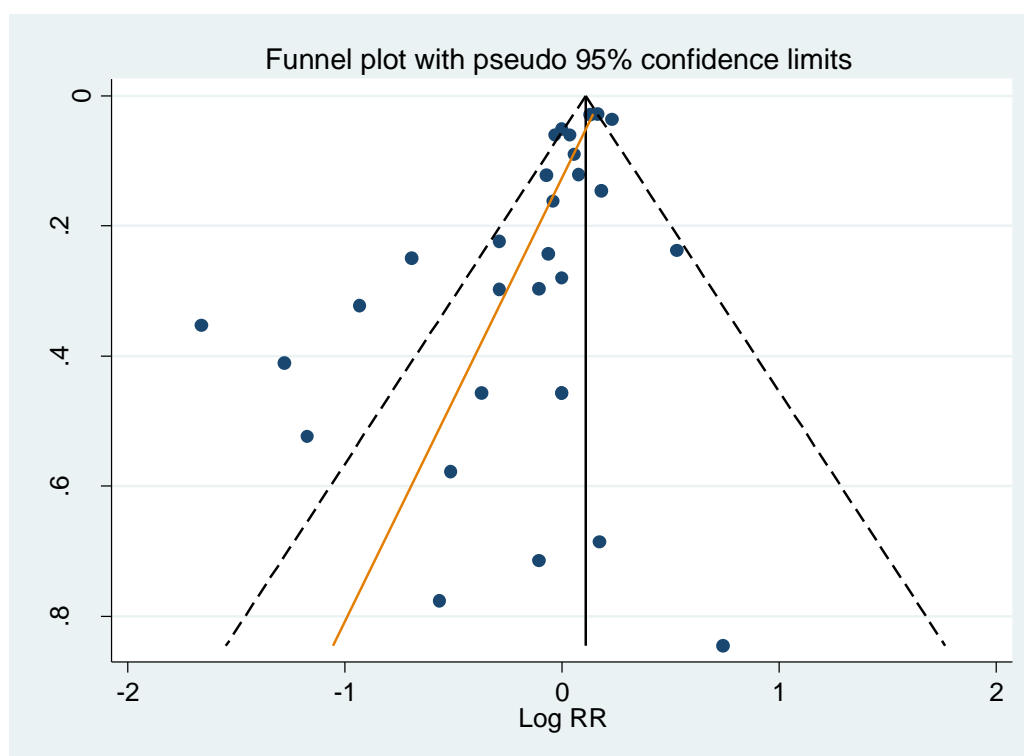
The most influential study in terms of altering the meta-RR when excluded was the study of pathologists by Hall *et al.* However, the overall findings were relatively robust to the exclusion of a single study.

In terms of the effect on the analysis, of possible alternative risk estimates being chosen, none of the studies with a weight of 5% or more in the main analysis had alternative

estimates and so this was judged not likely to have any impact on the overall conclusions.

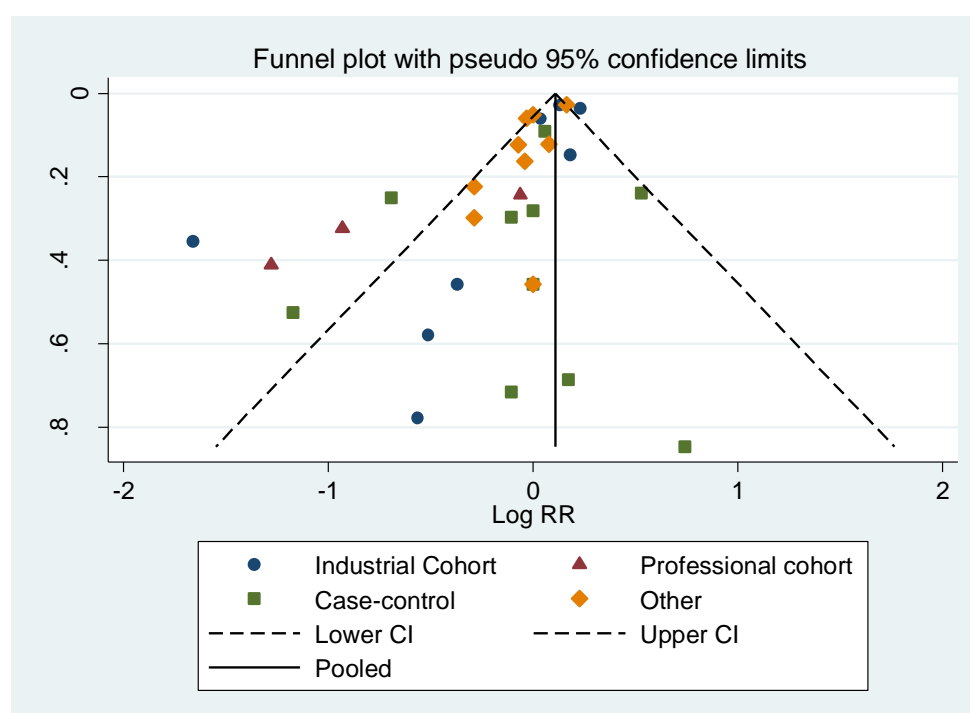
The existence of publication bias was first examined using a funnel plot. The plot contains the regression line corresponding to the regression test for funnel-plot symmetry proposed by Egger *et al* [26]. This is set out in Figure 5.4. There is a suggestion of asymmetry in the plot with a tendency for studies showing a lack of effect. However, a lack of symmetry in funnel plots does not necessarily indicate the presence of publication bias. The asymmetry in this plot is in the opposite direction to that expected with the usual publication bias (non-publication of small non-positive studies).

Figure 5.9 – Funnel plot of relative risk estimates for lung cancer in relation to occupational exposure to formaldehyde



Looking at the plot by study type yields the graph set out in Figure 5.10.

Figure 5.10 - Funnel plot of relative risk estimates for lung cancer in relation to occupational exposure to formaldehyde by study type



The estimated bias coefficient from Egger's test is -1.47 with a standard error of 0.38, giving a p-value of 0.001, which is highly statistically significant. Repeating the test excluding the professional cohort studies, yields a p-value of 0.004 which is still highly significant.

The significant excess of lung cancer from the fixed effect model was not evident in the random effects model, nor were there any statistically significant excesses in any of the important subgroups for this model.

The findings for lung cancer generally substantiate the prior belief that there was no association between occupational exposure to formaldehyde and increased risk of lung cancer, as exemplified by the IARC monograph not giving lung cancer a mention in its evidence synthesis statement [88]. The literature has recently focussed in on nasopharyngeal cancer and risk from haematopoietic cancers such as leukaemia, see for example [175]. The absence of an exposure-response relationship in studies also argued against a causal interpretation, although dose-response was not explicitly examined as

part of this analysis. The most recent meta-analyses that predated the current study that looked at lung cancer in relation to formaldehyde both found significant excess risks [76, 176]. Despite being chemically very reactive, it is plausible for formaldehyde to be adsorbed onto particles and be transported deep into the lungs [177]. However, there is no evidence that exposures received in industry are substantially higher than those in, for example, funeral homes.

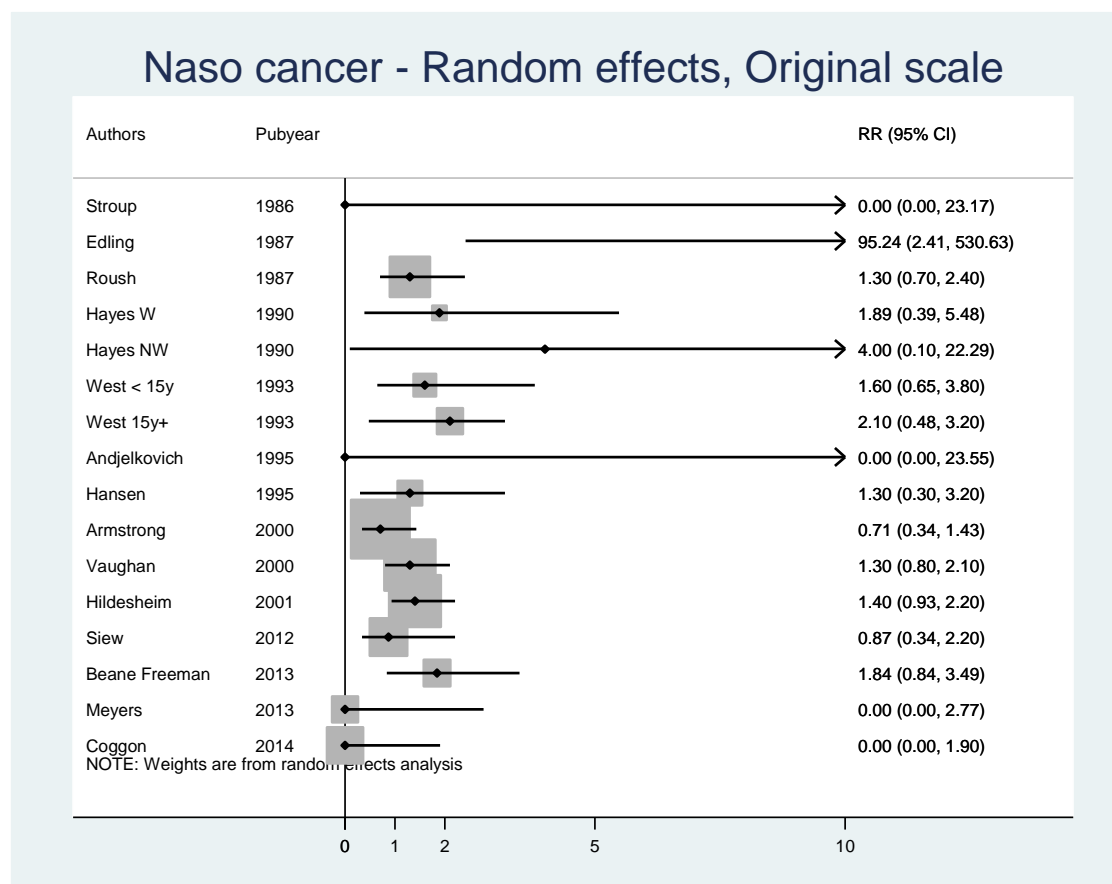
It is noteworthy that two of the large industrial cohort studies found a statistically significantly raised relative risk [89, 126]. It is plausible that this could be due to exposure to other occupational lung carcinogens likely to be present in the manufacturing plants included in the studies.

5.3.3 Nasopharyngeal cancer

For nasopharyngeal cancer there were 16 risk estimates from 14 studies containing 240 nasopharyngeal cancers. The fixed effect analysis on the original scale gave a meta-RR of 1.19 (0.94 to 1.44). Between-study heterogeneity was highly statistically significant and so the random effects analysis is to be preferred. Here the meta-RR was 1.41 (0.60 to 2.22). It is noteworthy that the analysis on the original scale allowed inclusion of four relative risk estimates where there were no observed cases.

The forest plot for all risk estimates is set out in Figure 5.11

Figure 5.11 – Random effects meta-analysis of nasopharyngeal cancer on the original scale



Note: the meta-RR was 1.41 (0.60 to 2.22). In order to avoid negative lower confidence limits a version of metan that respected CIs is given but this meant the meta-RR was excluded.

A sensitivity analysis excluding each study in turn from the fixed effect analysis is set out in Table 5.12.

Table 5.12 – Meta-RR for fixed effect model on original scale excluding each study in turn

Study Excluded	Meta-RR	95% CI
	1.41	0.60 to 2.22`
Edling	1.17	0.92 to 1.43
Roush	1.44	0.55 to 2.33
Hayes	1.29	0.43 to 2.15
West	1.36	0.46 to 2.25
Hildesheim	1.45	0.52 to 2.37
Armstrong	1.52	0.59 to 2.45
Hansen	1.43	0.57 to 2.28
Vaughan	1.45	0.53 to 2.37
Beane Freeman	1.38	0.51 to 2.26
Siew	1.48	0.59 to 2.37
Andjelkovich	1.44	0.62 to 2.25
Stroup	1.44	0.62 to 2.25
Coggon	1.50	0.66 to 2.34
Meyers	1.51	0.67 to 2.36

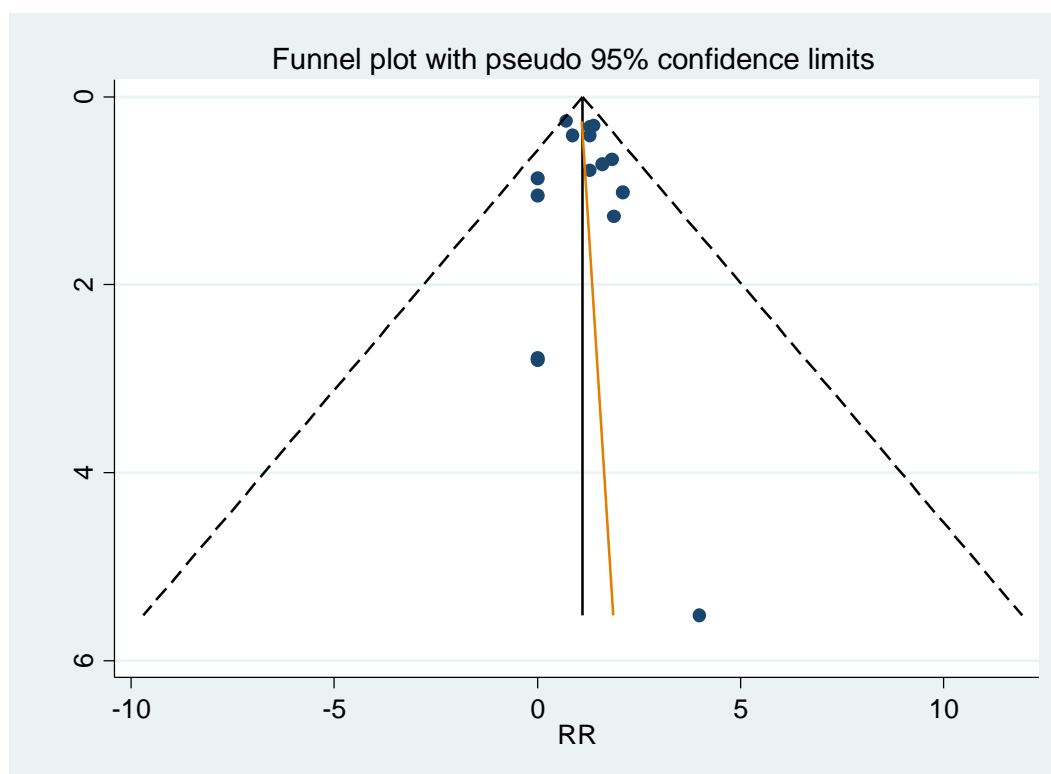
Of these studies, the exclusion of the Edling study has the largest influence on the meta-RR as it decreased to 1.17 although it remained not statistically significant.

Formal sub-group analyses would of limited value given the Edling outlier and the sensitivity of pooled estimate to method.

The funnel plot of relative risk estimates for nasopharyngeal cancer in relation to occupational exposure to formaldehyde is set out in Figure 5.12.

Note that this plot excludes the RR from the Edling study because of its very large standard error. There is a lack of asymmetry and the Egger regression test is not significant.

Figure 5.12 – Funnel plot of relative risk estimates for nasopharyngeal cancer in relation to occupational exposure to formaldehyde



The findings of this meta-analysis are consistent with the most recent meta-analyses conducted by Bachand *et al* [122], Bosetti *et al* [121], and Collins *et al* [76], suggesting that there is insufficient evidence of a statistically significantly elevated meta-RR. It may turn out that formaldehyde does cause nasopharyngeal cancer, but on the present evidence, there is insufficient evidence of a raised relative risk. This is further exacerbated by the relatively poor exposure assessment from the case-control studies and the problems in the interpretation of the excess found in one plant in the NCI cohort, albeit with the latter being a post hoc finding.

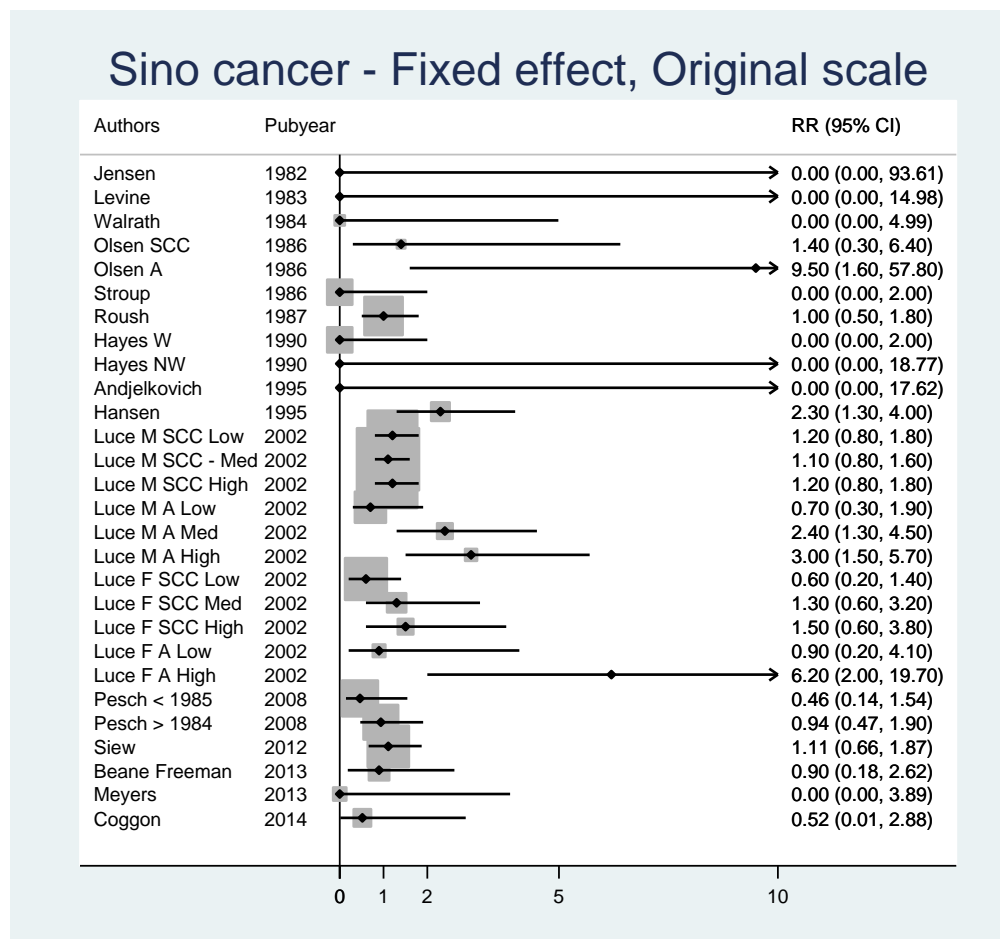
5.3.4 Sinonasal cancer

For sinonasal cancer there were 28 estimates from 15 studies incorporating 367 cases. The fixed effect analysis on the original scale gave a meta-RR of 1.01 (0.85 to 1.17). Between-study heterogeneity was not statistically significant and so the fixed effect

analysis is to be preferred. It is noteworthy that the analysis on the original scale allowed inclusion of eight relative risk estimates where there were no observed cases.

The forest plot for all risk estimates is set out in Figure 5.13.

Figure 5.13 – Fixed effect meta-analysis of sinonasal cancer on the original scale



Note: the meta-RR was 1.01 (0.85 to 1.17). In order to avoid negative lower confidence limits a version of metan that respected CIs is given but this meant the meta-RR was excluded.

A sensitivity analysis excluding each study in turn from the fixed effect analysis is set out in Table 5.14.

Table 5.14 – Meta-RR for fixed effect model on original scale excluding each study in turn

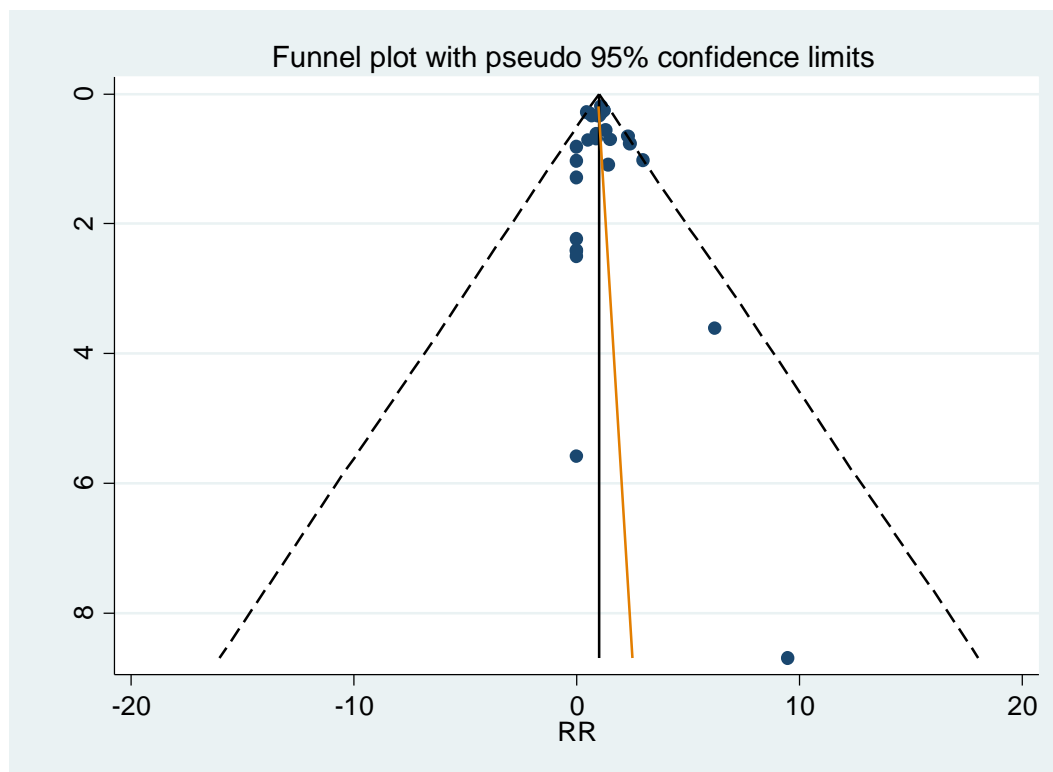
Study Excluded	Meta-RR	95% CI
	1.01	0.85 to 1.17
Olsen	1.00	0.85 to 1.17
Roush	1.01	0.85 to 1.18
Hansen	0.98	0.82 to 1.15
Beane Freeman	1.01	0.85 to 1.17
Coggon	1.01	0.85 to 1.18
Luce	0.86	0.59 to 1.13
Siew	1.00	0.83 to 1.17
Pesch	1.07	0.89 to 1.24
Andjelkovich	1.01	0.85 to 1.17
Stroup	1.01	0.85 to 1.17
Jensen	1.01	0.85 to 1.17
Hayes	1.02	0.85 to 1.18
Walrath	1.01	0.85 to 1.17
Levine	1.01	0.85 to 1.17
Meyers	1.01	0.85 to 1.18

Of these studies, the exclusion of the Luce study, not surprisingly, has the largest influence on the meta-RR as it reduces to 0.86 and therefore not statistically significantly raised.

Exploration of subgroup meta-RRs is not presented for sinonasal cancer, because of the absence of any evidence of heterogeneity, but tests of variation of subgroups were carried out and no significant differences were found.

The funnel plot of relative risk estimates for nasopharyngeal cancer in relation to occupational exposure to formaldehyde is set out in Figure 5.14.

Figure 5.14 – Funnel plot of relative risk estimates for sinonasal cancer in relation to occupational exposure to formaldehyde



There is a lack of asymmetry and the Egger regression test is not significant.

The findings of this meta-analysis are consistent with the most recent meta-analyses conducted by Binazzi *et al*, [123], Bachand *et al* [122], and Bosetti *et al* [121], suggesting that there is insufficient evidence of a statistically significantly elevated meta-RR.

5.4 DISCUSSION AND CONCLUSIONS

This meta-analysis substantiates the available evidence on risks of lung cancer and sinonasal cancer from occupational exposure to formaldehyde. The meta-RR for lung cancer, based on the random effects meta-analysis on the log scale was 0.98 (0.90 to

1.07) suggesting no overall increase in relative risk. For sinonasal cancer, the fixed effect meta-analysis gave a meta-RR of 1.01 (0.85 to 1.17).

The random effects meta-analysis for nasopharyngeal cancer gave a meta-RR of 1.41 (0.60 to 2.22). A naïve fixed effect meta-analysis on the log scale (thus excluding the studies with no observed cases) gave a meta-RR of 1.38 (1.11 to 1.71), a meta-relative risk which was of similar magnitude, but more precise. Examination of the totality of the evidence, including studies that have zero observed cases, does not provide support for the IARC classification that formaldehyde is a definite carcinogen in relation to nasopharyngeal cancer.

An association has been postulated for lung cancer on the basis that formaldehyde can be adsorbed onto the surfaces particles and therefore taken into the inner lung. The association with sinonasal cancer was considered as the nose and nasal cavities are the first point of contact for formaldehyde when breathed in through the nose.

Nasopharyngeal cancer however, is thought to be the most plausible respiratory cancer site, because it is accepted that formaldehyde causes sinonasal cancer in rates and the equivalent organ site in humans is the nasopharynx [96].

For lung cancer, significant heterogeneity was explored via subgroup analysis using the random effects model. However, none of the subgroups showed a statistically significant excess. This finding was robust to sensitivity analyses such as excluding one study at a time from the analysis and from choosing alternative risk estimates and are generally consistent with all the previously conducted meta-analyses [76, 117-119, 122, 124, 176]. Likewise, the findings for sinonasal cancer and nasopharyngeal cancer were also consistent with the earlier meta-analyses.

Some aspects of study heterogeneity, in particular, differences in exposure levels and methods of exposure assessment were not explored in detail in these analyses and

should be examined, as they have the potential to mask differences between study types. Some of the studies were also limited, in some cases by design and in others by conduct, in their ability to control for known and suspected confounding factors, either occupational or non-occupational in origin. However, it is noteworthy that whether or not a study adjusted for smoking did not seem to have any impact on the meta-analyses conducted.

The meta-analyses carried out here used an indirect method of calculating an expected number of cases for the rarer cancers based on an appropriate fraction of the expected number of lung cancers. Other approaches to deriving an expected number are possible and it would be a useful comparison exercise for this to be done. Only studies where an actual number of cases could be deduced from reading the manuscripts were included. Although efforts were made to include studies where an observed number of cancers could be deduced (and therefore a relative risk could be estimated), studies where this could not be done were not included. This has not been attempted, but might further reduce the meta-RR for both nasopharyngeal and sinonasal cancers were such an exercise to be carried out.

A limitation of the meta-analyses carried out here is that some of the studies were based on administrative data, such as cancer registries, and thus often lacked information on risk factors and estimates of formaldehyde exposure. Indirect methods of assessing exposure to formaldehyde were used based on job history, rather than on actual measurement data.

Other aspects importantly not controlled for in the meta-analyses were differences in the baseline characteristics of the study populations and differences in length of follow-up for cohort studies making the pooling of SMRs questionable.

The lack of heterogeneity in the sinonasal cancer analyses could be due to a lack of statistical power to detect it.

Future research should aim to collect more detailed data on study characteristics and should assess the exposure-response relationship, perhaps by indirectly assessing exposures for those studies lacking in an exposure assessment. Assessing the potential for an impact of co-exposures should also be considered.

6 CONCLUSIONS

6.1 OVERVIEW

This thesis set out with the intention of improving on current practice of meta-analyses of rare diseases in occupational epidemiology. The example which motivated the work was prompted by the announcement by the International Agency for Research on Cancer that formaldehyde was a definite carcinogen in humans, because it was a proven carcinogen for sinonasal cancer in rats and from occupational cohort and population-based case-control studies of nasopharyngeal cancer in humans [178]. The evidence from occupational cohort studies seemed over-reliant on the single study carried out by the National Cancer Institute in the USA [82]. A potentially important source of bias for these studies could have been the exclusion of studies with no exposed cases. The evidence was also reliant on a number of population-based case-control studies that can be prone to bias [179]. This could have arisen in two important ways: inadequate control of exposure to Epstein Barr virus, a known risk factor for nasopharyngeal cancer, the prevalence of which is particularly high in Asian populations could have resulted in a selection bias; poor characterisation and control of known or suspected confounding factors.

Eight meta-analyses of one or more of lung cancer, nasopharyngeal cancer and sinonasal cancer have been published in the epidemiological literature [76, 117, 118, 121-123, 180, 181]. None of them has found strong evidence of an increased risk for any of these cancers, making the most recent IARC findings [88] somewhat at odds with the totality of the literature.

6.2 PAST PRACTICE OF META-ANALYSIS IN OCCUPATIONAL EPIDEMIOLOGY

The published review of past practice only contained a review of practice until 2001 and so is now somewhat out of date. The call for better approaches to integrate sources of variation including potentially diverse measures of occupational exposure and SMRs based on different study populations, comparator populations and lengths of follow remains as relevant today as it did then. Petitti's cautionary note on meta-analysis [2] is better heeded now than in was some 10 years ago. Reporting of meta-analyses according to guidelines such as the MOOSE guidelines [128] are generally insisted upon by journal editors of the better occupational epidemiology journals. They importantly recommend including a flow diagram of study identification and selection, double independent data extraction, robust classification of exposure to the agent or substance of interest as well as any important known or suspected occupational and non-occupational confounding factors and exploration of heterogeneity (now becoming known as bias exploration in some fields [182]). Appropriate sensitivity analysis to test the robustness of any findings such as in Bachand *et al* [122] should also be included.

Therefore, although current practice has certainly improved since the earlier review [3], there remains many unanswered issues about the robustness of the meta-analysis methodology as employed in occupational epidemiology, particularly with respect to rare diseases such as nasopharyngeal cancer and sinonasal cancer. In particular the issues of diverse measures of occupational exposures between studies and the comparability and combination of SMRs have not yet been fully resolved.

6.3 REVIEW OF METHODS

Specific methods exist for combining odds ratios, SMRs, correlation coefficients and generic measures of relative risk. A fixed effect approach exists for estimating the

meta-RR in the absence of between-study heterogeneity exists as does a random effects approach for the presence of significant heterogeneity. Bayesian methods also exist. Heterogeneity can be explored informally via subgroup analysis or via regression modelling, either via meta-regression or for observed cases from cohort studies via Poisson or negative binomial regression. These models can also incorporate random effects terms. Methods also exist for assessing and testing for publication bias. Specific methods for dealing with exposure-response data also exist, but they are outside the scope of this research. Sensitivity analyses to assess the robustness of any findings now play a key role in any meta-analysis. More recent developments include, particularly in multivariate method analysis and in network analysis. A very recent development is the so-called IVhet model for which it is claimed to be an improvement on both fixed effect and random effects models.

6.4 RARE DISEASE OUTCOMES

The use of adding a small correction to the observed number of cases or the observed and expected number of cases seems to be problematic as the optimum value of this correction, that is the one that minimizes the introduction of bias into the analysis, cannot be chosen in advance, but depends on the data being analysed. Methods that utilise analysis on the log scale and by so doing exclude studies are to be avoided because of the inflationary tendency for the relative risks in such studies. Analyses on the original untransformed scale obviate the need to exclude such studies. However, further exploration of the reliance of this approach on a normality assumption is warranted. Further exploration of Poisson regression and in the presence of over-dispersion negative binomial fixed and random effects models for their relevance for occupational epidemiological studies should be undertaken (there may be difficulties for studies other than cohort studies).

6.5 META-ANALYSES OF FORMALDEHYDE

The most recent meta-analysis of formaldehyde and lung cancer gave a meta-RR of 1.07 (1.01 to 1.12) for industrial cohort studies and 0.68 for professional cohort studies [121]. The most recent meta-analysis of lung cancer in case-control studies yielded a meta-RR of 0.8 (0.7 to 1.0) [76]. For nasopharyngeal cancer the meta-RR for industrial cohort studies was 0.72 (0.40 to 1.29) [122], for professional cohort studies was 0.91 (0.25 to 2.33) [76] and for case-control studies was 1.22 (1.00 to 1.50) [122]. For sinonasal cancer the meta-RR for industrial cohort studies was 1.09 (0.66 to 1.79) [123] for professional cohort studies was 0.00 (0.00 to 3.51) [183] and for case-control studies was 1.68 (1.37 to 2.06) [123].

For lung there were 30 estimates of relative risk incorporating 6,476 lung cancers. The fixed effect analysis gave a meta-RR of 1.12 (1.08 to 1.15). However, there was significant heterogeneity ($p < 0.001$). The random effects model gave a meta-RR of 0.98 (0.90 to 1.07) suggesting that there was no association between occupational exposure to formaldehyde and increased risk of lung cancer. This finding was consistent with the earlier meta-analyses.

None of the subgroups defined a priori contained a statistically significant excess. The findings were robust to the exclusion of single studies and also to the choice of alternative (suboptimal) effect estimates. There was no evidence of publication bias in the form of likely missing small studies showing small effects. This finding was consistent with previous meta-analyses and also with the views of the IARC working group.

There were a number of limitations to the lung cancer analysis, the most important of which was the lack of consideration of exposure-response. Exposure-response is considered an important aspect of the Bradford Hill causality considerations [184].

Good practice would normally involve at least two researchers independently extracting the data for the study and resolving any differences found. This did not happen and would need to happen before the substantive meta-analysis could be written up for submission for publication in a peer-reviewed scientific journal. Other forms of bias assessment could also be carried out in relation to the assessment of ever or never exposed to formaldehyde. It is likely that this assessment might vary over time for example, and it is a moot point as to what is regarded as unexposed for a substance that is regarded as ubiquitous in the environment. This is probably best undertaken as part of consideration of exposure response. A consistent assessment of exposure across the studies in the absence of knowledge of study outcome is required and is probably best carried out by an experienced occupational hygienist or exposure scientist. Other sources of heterogeneity include the inclusion and exclusion criteria applied to study populations, although these are not always reported, together with the length of follow-up in occupational cohort studies. Ideally an adjustment for individual-level exposure to tobacco smoking and other known or suspected occupational lung carcinogens would be available, but this rarely happens in practice. Nevertheless, some indication of the potential impact on the meta-RR would be useful. Thus although the confidence intervals from the random effects meta-analysis reflect statistical heterogeneity, other variations between studies suggest that the true level of uncertainty surrounding the meta-RR is probably greater than that expressed by the confidence interval.

It is noteworthy that two of the large industrial cohort studies found a statistically significantly raised relative risk [89, 126]. It is plausible that this could be due to exposure to other occupational lung carcinogens likely to be present in the manufacturing plants included in the studies.

For nasopharyngeal cancer there were 16 risk estimates from 14 studies containing 240 nasopharyngeal cancers. Four studies contained no observed cases. The random effects

meta-analysis on the original scale gave a meta-RR of 1.41 (0.60 to 2.22). This finding is consistent with previous meta-analyses, for example Bachand *et al*, Bosetti *et al* and Collins *et al* [122, 176, 185] that have been carried out, but does not provide strong support for the IARC view that exposure to formaldehyde causes nasopharyngeal cancer [88].

Formaldehyde causes nasal cancer in rats and due to differences in the nature of the nasal passage between rats and humans; the equivalent site in humans is the nasopharynx. However, the meta-analyses undertaken here is not supportive of a casual interpretation. It may turn out with further evidence that formaldehyde causes nasopharyngeal cancer, but the evidence to support this conclusion is insufficient. The same shortcomings that apply to the evidence relating to lung cancer also apply to nasopharyngeal cancer. In addition, the inclusion of evidence from studies that had the potential to report on nasopharyngeal cancer, but provided no way of deducing the number of cases, could be additionally included in the meta-analysis assuming there were no actual cases. Alternative methods of estimating the expected number of nasopharyngeal cancer cases, rather than taking a fixed fraction of the expected number of lung cancer cases might also be explored. Whilst the argument in the literature over plant 1 in the NCI cohort study continues [91], it appears to not unduly influence the findings of the meta-analysis presented in this thesis as the study has much less influence on the meta-RR than for example the Armstrong case-control study [98]. There appears to be no evidence of publication bias for nasopharyngeal cancer in relation to formaldehyde exposure.

For sinonasal cancer there were 28 estimates from 15 studies incorporating 367 cases. The fixed effect analysis on the original scale gave a meta-RR of 1.01 (0.89 to 1.17). There was no significant between-study heterogeneity and the findings were robust to the exclusion of individual studies and no evidence of publication bias. A similar

discussion applies to that for nasopharyngeal cancer, except that the nasopharynx is the more relevant site of contact for formaldehyde than the nose and nasal cavities. It is also worth noting that an additional limitation of the sinonasal cancer analysis is the lack of separate (due to small numbers) of the nasal cancers from those of the sinonasal cavities as it is possible that the two cancer sites may have different aetiologies.

6.6 CLOSING REMARKS

Meta-analysis has been used to demonstrate that there is a lack of evidence that occupational exposure to formaldehyde increases the risk of nasopharyngeal cancer, sinonasal cancer or lung cancer. Studies of rare diseases in occupational epidemiology should take care to include evidence from studies with no observed cases as the resulting bias could lead to spuriously significant excesses being found.

For future research priorities, my experience suggests that bias analysis methods should be exploited as there is a significant potential for bias in studies in occupational epidemiology. The combination of risk estimates from different study types and whether or not that is appropriate still requires resolution. More important for occupational epidemiology is whether SMRs can be combined to produce a meta-SMR when heterogeneous study populations and comparison populations are used and cohorts have different lengths of follow-up and exposures in the cohorts may not be directly comparable. Some clear guidance on whether and when relative risk estimates from different study types might be combined would be useful. Meta-analysis seems to be used a little more cautiously now in occupational epidemiology. Generally studies follow the MOOSE reporting guidelines and the robustness of findings and assessment of publication bias are carried out. However, other forms of bias, principally around study characteristics and exposure data require resolution. It is important this issue is brought to the attention of research funders, as at present, most of the methodological

development in meta-analyses occurs within the randomized controlled clinical trial setting. However, the issues specific to occupational epidemiology may not be being addressed as a matter of priority. There is a growing awareness of the issue of meta-analysis of rare diseases in occupational epidemiology and the methodology for dealing with is partly stems from that developed for dealing with rare adverse events in meta-analyses of randomized controlled clinical trials. However there remains a current lack of consensus on how this issue should be dealt with. Finally, the statistical issues relating to analysis on the original scale require clarifying as to whether distributional assumptions have been violated.

Meta-analysis remains a powerful technique for helping to make sense of associations in a variety of occupational and population settings. However, a number of unresolved issues remain with its application in the field of occupational epidemiology.

REFERENCES

1. Glass GV. Primary, secondary and meta-analysis of research. Educational Researcher 1976; 5: 3-8.
2. Petitti DB. Of babies and bathwater. American Journal Of Epidemiology 1994; 140: 779-782.
3. McElvenny DM, G AB, Jarup L, Higgins JP. Meta-analysis in occupational epidemiology: a review of practice. Occupational Medicine (Oxford, England) 2004; 54: 336-344.
4. Verbeek H, Ijaz S. Systematic reviews of occupational health and safety topics. In Venables K (ed) Current Topics in Occupational Epidemiology, Edition Oxford: Oxford University Press 2013.
5. Rodgers M, Sowden A, Petticrew M et al. Testing Methodological Guidance on the Conduct of Narrative Synthesis in Systematic Reviews: Effectiveness of Interventions to Promote Smoke Alarm Ownership and Function. Evaluation 2009; 15: 49-73.
6. Garg AX, Hackam D, Tonelli M. Systematic Review and Meta-analysis: When One Study Is Just not Enough. Clinical Journal of the American Society of Nephrology 2008; 3: 253-260.
7. Fleiss JL. The statistical basis of meta-analysis. Statistical Methods In Medical Research 1993; 2: 121-145.
8. Moulin JJ. A meta-analysis of epidemiologic studies of lung cancer in welders. Scandinavian Journal Of Work, Environment & Health 1997; 23: 104-113.
9. Frumkin H, Berlin J. Asbestos exposure and gastrointestinal malignancy review and meta-analysis. American Journal Of Industrial Medicine 1988; 14: 79-95.

10. Herbold M. Meta-analysis of environmental and occupational epidemiological studies: a method demonstrated using the carcinogenicity of PCBs as an example. *Sozial-Und Praventivmedizin* 1993; 38: 185-189.
11. Mantel N, Haenszel W. Statistical aspects of the analysis of data from retrospective studies of disease. *Journal of the National Cancer Institute* 1959; 22: 719-723.
12. Robins J, Greenland S, Breslow NE. A general estimator for the variance of the Mantel-Haenszel odds ratio. *American Journal of Epidemiology* 1986; 124: 719-723.
13. Robins J, Breslow N, Greenland S. Estimators of the Mantel-Haenszel variance consistent in both sparse data and large-strata limiting models. *Biometrics* 1986; 42: 311-323.
14. Greenland S, Salvan A. Bias in the one-step method for pooling study results. *Statistics In Medicine* 1990; 9: 247-252.
15. Yusuf S, Peto R, Lewis J et al. Beta blockade during and after myocardial infarction: an overview of the randomized trials. *Progress in Cardiovascular Diseases* 1985; 27: 335-371.
16. Hedges LV, Olkin I. Vote-counting methods in research synthesis. *Psychological Bulletin* 1980; 88: 359-369.
17. Kelley GA, Kelley KS. Statistical models for meta-analysis: A brief tutorial. *World Journal of Methodology* 2012; 2: 27-32.
18. DerSimonian R, Laird N. Meta-analysis in clinical trials. *Controlled Clinical Trials* 1986; 7: 177-188.
19. Thompson SG, Sharp SJ. Explaining heterogeneity in meta-analysis: a comparison of methods. *Statistics In Medicine* 1999; 18: 2693-2708.

20. Emerson JD. Combining estimates of the odds ratio: the state of the art. *Statistical Methods In Medical Research* 1994; 3: 157-178.
21. Spiegelhalter DJ, Thomas A, Best N, Lunn D. Winbugs Version 1.4 User Manual. Cambridge: MRC Biostatistics Unit 2003.
22. Sutton AJ, Abrams KR. Bayesian methods in meta-analysis and evidence synthesis. *Statistical Methods In Medical Research* 2001; 10: 277-303.
23. Hardy RJ, Thompson SG. Detecting and describing heterogeneity in meta-analysis. *Statistics in Medicine* 1998; 17: 841-856.
24. Greenland S. Quantitative methods in the review of epidemiologic literature. *Epidemiologic Reviews* 1987; 9: 1-30.
25. Galbraith RF. A note on graphical presentation of estimated odds ratios from several clinical trials. 1988.
26. Egger M, Davey Smith G, Schneider M, Minder C. Bias in meta-analysis detected by a simple, graphical test. *BMJ (Clinical Research Ed.)* 1997; 315: 629-634.
27. Begg CB, Mazumdar M. Operating characteristics of a rank correlation test for publication bias. 1994.
28. Duval S, Tweedie R. Trim and fill: A simple funnel-plot-based method of testing and adjusting for publication bias in meta-analysis. *Biometrics* 2000; 56: 455-463.
29. Rosenthal R. The 'file drawer problem' and tolerance of null results. *Psychological Bulletin* 1979; 86: 638-641.
30. Gleser LJ, Olkin I. Models for estimating the number of unpublished studies. *Statistics in Medicine* 1996; 15: 2493-2507.
31. Copas J. What works? Selectivity models and meta-analysis. *J R Statist Soc A* 1999; 162: 177-188.

32. Greenland S, Longnecker MP. Methods for trend estimation from summarized dose-response data, with applications to meta-analysis. *American Journal Of Epidemiology* 1992; 135: 1301-1309.
33. Shadish WR, Haddock CK. Combining estimates of effect size. In Cooper H, Hedges LV (eds): *Handbook of Research Synthesis*, Edition New York: Russel Sage Foundation 1994.
34. Greenland S, O'rourke K. On the bias produced by quality scores in meta-analysis, and a hierarchical view of proposed solutions. *Biostatistics* 2001; 2: 463-471.
35. Wong O, Raabe GK. Critical review of cancer epidemiology in petroleum industry employees, with a quantitative meta-analysis by cancer site. *American Journal Of Industrial Medicine* 1989; 15: 283-310.
36. Weiss W. Asbestos and colorectal cancer. *Gastroenterology* 1990; 99: 876-884.
37. Rothman KJ. Cancer occurrence among workers exposed to acrylonitrile. *Scandinavian Journal Of Work, Environment & Health* 1994; 20: 313-321.
38. Khuder SA, Mutgi AB. Meta-analyses of multiple myeloma and farming. *American Journal Of Industrial Medicine* 1997; 32: 510-516.
39. Blair A, Zahm SH, Pearce NE et al. Clues to cancer etiology from studies of farmers. *Scandinavian Journal Of Work, Environment & Health* 1992; 18: 209-215.
40. Kheifets LI, Afifi AA, Buffler PA et al. Occupational electric and magnetic field exposure and leukemia. A meta-analysis. *Journal Of Occupational And Environmental Medicine / American College Of Occupational And Environmental Medicine* 1997; 39: 1074-1091.
41. Sutton AJ, Higgins JPT. Recent developments in meta-analysis. *Statistics in Medicine* 2008; 27: 625-650.

42. Reitsma JB, Glas AS, Rutjes AWS et al. Bivariate analysis of sensitivity and specificity produces informative summary measures in diagnostic reviews. *Journal of Clinical Epidemiology* 2012; 58: 982-990.
43. Jackson D, Riley R, White IR. Multivariate meta-analysis: Potential and promise. *Statistics in Medicine* 2011; 30: 2481-2498.
44. Siontis GCM, Stefanini GG, Mavridis D et al. Percutaneous coronary interventional strategies for treatment of in-stent restenosis: a network meta-analysis. *The Lancet* 2015; 386: 655-664.
45. Lumley T. Network meta-analysis for indirect treatment comparisons. *Statistics in Medicine* 2002; 21: 2313-2324.
46. Wei Y, Higgins JPT. Estimating within-study covariances in multivariate meta-analysis with multiple outcomes. *Statistics in Medicine* 2013; 32: 1191-1205.
47. Clark L, Fairhurst C, Hewitt CE et al. A methodological review of recent meta-analyses has found significant heterogeneity in age between randomized groups. *Journal of Clinical Epidemiology* 2014; 67: 1016-1024.
48. Doi SAR, Barendregt JJ, Khan S et al. Simulation Comparison of the Quality Effects and Random Effects Methods of Meta-analysis. *Epidemiology* 2015; 26: e42-e44.
49. Langan D, Higgins JPT, Simmonds M. Comparative performance of heterogeneity variance estimators in meta-analysis: a review of simulation studies. *Research Synthesis Methods* 2016; n/a-n/a.
50. Paule RC, Mandel J. Consensus values and weighting factors. *Journal of Research of the National Bureau of Standards* 1982; 87: 377-385.
51. Jørgensen L, Paludan-Müller AS, Laursen DRT et al. Evaluation of the Cochrane tool for assessing risk of bias in randomized clinical trials: overview

- of published comments and analysis of user practice in Cochrane and non-Cochrane reviews. *Systematic Reviews* 2016; 5: 80.
52. Spittal MJ, Pirkis J, Gurrin LC. Meta-analysis of incidence rate data in the presence of zero events. *BMC Medical Research Methodology* 2015; 15: 1-16.
 53. StataCorp. Stata Statistical Software: Release 13. In Edition College Stations, TX: StataCorp LP: 2013.
 54. Sankey SS, Weissfeld LA, Fine MJ, Kapoor W. An assessment of the use of the continuity correction for sparse data in metaanalysis. *Communications in Statistics-Simulation and Computation* 1996; 25: 1031-1056.
 55. Austin H, Perkins LL, Martin DO. Estimating a relative risk across sparse case-control and follow-up studies: a method for meta-analysis. *Statistics In Medicine* 1997; 16: 1005-1015.
 56. Sutton AJ, Cooper NJ, Lambert PC et al. Meta-analysis of rare and adverse event data. *Expert Rev Pharmacoeconomics Outcomes Res* 2002; 2: 367-379.
 57. Sweeting MJ, Sutton AJ, Lambert PC. What to add to nothing? Use and avoidance of continuity corrections in meta-analysis of sparse data. *Statistics In Medicine* 2004; 23: 1351-1375.
 58. Fleiss JL. Measures of effect size for categorical data. In Cooper H, Hedges LV (eds): *The handbook of research synthesis*, Edition New York: Russel Sage Foundation 1994; 245-260.
 59. Bradburn MJ, Deeks JJ, Berlin JA, Localio AR. Much ado about nothing: A comparison of the performance of meta-analytical methods with rare events. *Stat Med* 2007; 26: 53-77.
 60. Woolf B. On estimating the relation between blood group and disease. *Ann Hum Genet* 1955; 19: 251-253.

61. Greenland S, Robins JM. Estimation of a common effect parameter from sparse follow-up data. *Biometrics* 1985; 41: 55-68.
62. Mehta C, Patel NR, Gray R. On comparing an exact confidence interval for the common odds ratio in several 2x2 contingency tables. *Journal of the American Statistical Association* 1985; 80: 969-973.
63. Deeks JJ, Altman DG, Bradburn MJ. Statistical methods for examining heterogeneity and combining results from several studies in meta-analysis. In Eggar M, Smith GD, Altman DG (eds): *Systematic reviews in healthcare: meta-analysis in context*, Edition London: BMJ Publishing Group 2001; 285-312.
64. McEntegart DJ. Pooling in integrated safety databases. *Drug Information Journal* 2000; 34.
65. Lievre M, Cucherat M, Leizorovicz A. Pooling, meta-analysis and the evaluation of drug safety. *Current Controlled Trials in Cardiovascular Medicine* 2002; 3: 6.
66. Friedrich JO, Beyene J, Adhikari NK. Rosiglitazone: can meta-analysis accurately estimate excess cardiovascular risk given the available data? Re-analysis of randomized trials using various methodologic approaches. *BMC Research Notes* 2009; 2: 1-10.
67. Stijnen T, Hamza TH, Özdemir P. Random effects meta-analysis of event outcome in the framework of the generalized linear mixed model with applications in sparse data. *Statistics in Medicine* 2010; 29: 3046-3067.
68. Cai T, Parast L, Ryan L. Meta-Analysis For Rare Events. *Statistics in Medicine* 2010; 29: 2078-2089.
69. Böhning D, Mylona K, Kimber A. Meta-analysis of clinical trials with rare events. *Biometrical Journal* 2015; 57: 633-648.

70. Bai O, Chen M, Wang X. Bayesian Estimation and Testing in Random Effects Meta-analysis of Rare Binary Adverse Events. *Statistics in biopharmaceutical research* 2016; 8: 49-59.
71. Bhaumik DK, Amatya A, Normand S-L et al. Meta-Analysis of Rare Binary Adverse Event Data. *Journal of the American Statistical Association* 2012; 107: 555-567.
72. Cox DR. The continuity correction. *Biometrika* 1970; 57: 217-219.
73. Agresti A. An introduction to categorical data analysis. New York: Wiley 1996.
74. Ulm K. Simple method to calculate the confidence interval of a standardized mortality ratio (SMR). *Am. J. Epidemiol.* 1990; 131: 373-375.
75. Sharp S, Sterne J. Meta-analysis. *Stata Technical Bulletin* 1997; 38: 9-14.
76. Collins JJ, Acquavella JF, Esmen NA. An updated meta-analysis of formaldehyde exposure and upper respiratory tract cancers. *J Occup Environ Med* 1997; 39: 639-651.
77. Jansen M. Multiscale Poisson data smoothing. *Journal of the Royal Statistical Society. Series B.* 2006; 68: 27-48.
78. Oehlert GW. A note on the delta method. *The American Statistician* 1992; 46: 27-29.
79. Seber GUH. Poisson regression. In Gail MH, Benichou J (eds): *Encyclopedia of Epidemiologic Methods*, Edition Chichester: John Wiley & Sons Ltd 1999.
80. Martuzzi M, Hills M. Estimating the Degree of Heterogeneity between Event Rates Using Likelihood. *Am. J. Epidemiol.* 1995; 141: 369-374.
81. Rothman KJ, Boice JD. *Epidemiological analysis with a programmable calculator*. Washington, D.C.: US Government Printing Office 1979.

82. Hauptmann M, Lubin JH, Stewart PA et al. Mortality from solid cancers among workers in formaldehyde industries. *American Journal Of Epidemiology* 2004; 159: 1117-1130.
83. Marsh GM, Youk AO, Buchanich JM et al. Pharyngeal cancer mortality among chemical plant workers exposed to formaldehyde. *Toxicology And Industrial Health* 2002; 18: 257-268.
84. Marsh GM, Youk AO. Reevaluation of mortality risks from nasopharyngeal cancer in the formaldehyde cohort study of the National Cancer Institute. *Regul Toxicol Pharmacol* 2005; 42: 275-283.
85. Cogliano VJ, Grosse J, Bann RA et al. Meeting report: summary of IARC monographs on formaldehyde, 2-butoxyethanol and 1-tert-butoxy-2-propanol. *Env Health Perspective* 2005; 113: 1205-1208.
86. Marsh GM, Youk AO, Buchanich JM et al. Work in the metal industry and nasopharyngeal cancer mortality among formaldehyde-exposed workers. *Regulatory Toxicology and Pharmacology* 2007; 48: 308-319.
87. Marsh GM, Youk AO, Morfeld P et al. Incomplete follow-up in the National Cancer Institute's formaldehyde worker study and the impact on subsequent reanalyses and causal evaluations. *Regulatory Toxicology and Pharmacology* 2010; 58: 233-236.
88. IARC. Chemical agents and related occupations. Lyon: IARC 2012.
89. Beane Freeman LE, Blair A, Lubin JH et al. Mortality from solid tumors among workers in formaldehyde industries: An update of the NCI cohort. *American Journal of Industrial Medicine* 2013; 56: 1015-1026.
90. Marsh GM, Morfeld P, Collins JJ, Symons JM. Issues of methods and interpretation in the National Cancer Institute formaldehyde cohort study. *Journal of Occupational Medicine and Toxicology* 2014; 9: 1-9.

91. Marsh GM, Morfeld P, Zimmerman SD et al. An updated re-analysis of the mortality risk from nasopharyngeal cancer in the National Cancer Institute formaldehyde worker cohort study. *Journal of Occupational Medicine and Toxicology* (London, England) 2016; 11: 8.
92. IARC. IARC Monographs on the Evaluation of Carcinogenic Risks to Humans: Volume 62. Wood dust and formaldehyde. Lyon, France: International Agency for Research on Cancer 1995.
93. Checkoway H, Pearce N, Dement JM. Design and conduct of occupational epidemiology studies: I. Design aspects of cohort studies. *American Journal of Industrial Medicine* 1989; 15: 363-373.
94. Squire RA, Cameron LL. An analysis of potential carcinogenic risk from formaldehyde. *Regulatory Toxicology And Pharmacology: RTP* 1984; 4: 107-129.
95. Conolly RB, Kimbell JS, Janszen DB, Miller FJ. Dose response for formaldehyde-induced cytotoxicity in the human respiratory tract. *Regulatory Toxicology And Pharmacology: RTP* 2002; 35: 32-43.
96. Kerns WD, Pavkov KL, Donofrio DJ et al. Carcinogenicity of formaldehyde in rats and mice after long-term inhalation exposure. *Cancer Research* 1983; 43: 4382-4392.
97. Coleman MP, Babb P, Damiecki P et al. Cancer survival trends in England and Wales, 1971-1995: deprivation and NHS region. London: The Stationery Office 1999.
98. Armstrong RW, Imrey PB, Lye MS et al. Nasopharyngeal carcinoma in Malaysian Chinese: occupational exposures to particles, formaldehyde and heat. *International Journal Of Epidemiology* 2000; 29: 991-998.

99. Siemiatycki J, Richardson L, Straif K et al. Listing occupational carcinogens. *Environmental Health Perspectives* 2004; 112: 1447-1459.
100. Mansfield CT, Hodge BT, Hege RB, Jr, Hamlin WC. Analysis of formaldehyde in tobacco smoke by high performance liquid chromatography. *Journal Of Chromatographic Science* 1977; 15: 301-302.
101. Stewart PA, Cubit D, Blair A. Formaldehyde levels in seven industries. *Appl Ind Hyg* 1987; 2: 231-236.
102. Kerfoot EJ, Mooney TF. Formaldehyde and paraformaldehyde study in funeral homes. *American Industrial Hygiene Association Journal* 1975; 36: 533-537.
103. Korczynski RE. Formaldehyde exposure in the funeral industry. *Appl Occup Environ Hyg* 1994; 9: 575-579.
104. Williams TM, Levine RJ, Blunden PB. Exposure of embalmers to formaldehyde and other chemicals. *American Industrial Hygiene Association Journal* 1984; 45: 172-176.
105. Stewart PA, Herrick RF, Feigley CE et al. Study design for assessing exposure of embalmers for a case-control study. Part 1. Monitoring results. *Appl Occup Environ Hyg* 1992; 7: 532-540.
106. Skisak CM. Formaldehyde vapor exposures in anatomy laboratories. *American Industrial Hygiene Association Journal* 1983; 44: 948-950.
107. Akbar-Khanzadeh F, Vaquerano MU, Akbar-Khanzadeh M, Bisesi MS. Formaldehyde exposure, acute pulmonary response, and exposure control options in a gross anatomy laboratory. *American Journal Of Industrial Medicine* 1994; 26: 61-75.
108. Perkins JL, Kimbrough JD. Formaldehyde exposure in a gross anatomy laboratory. *Journal Of Occupational Medicine.: Official Publication Of The Industrial Medical Association* 1985; 27: 813-815.

109. Coldiron VR, Ward JB, Jr, Trieff NM et al. Occupational exposure to formaldehyde in a medical center autopsy service. *Journal Of Occupational Medicine.: Official Publication Of The Industrial Medical Association* 1983; 25: 544-548.
110. Bernstein RS, Stayner LT, Elliott LJ et al. Inhalation exposure to formaldehyde: an overview of its toxicology, epidemiology, monitoring, and control. *American Industrial Hygiene Association Journal* 1984; 45: 778-785.
111. Luker MA, Van Houten RW. Control of formaldehyde in a garment sewing plant. *Am Ind Hyg Assoc J* 1990; 51: 541-544.
112. Priha E, Riipinen H, Korhonen K. Exposure to formaldehyde and solvents in Finnish furniture factories in 1975-1984. *The Annals Of Occupational Hygiene* 1986; 30: 289-294.
113. Sass-Kortsak AM, Holness DL, Pilger CW, Nethercott JR. Wood dust and formaldehyde exposures in the cabinet-making industry. *American Industrial Hygiene Association Journal* 1986; 47: 747-753.
114. Partanen T, Kauppinen T, Nurminen M et al. Formaldehyde exposure and respiratory and related cancers. A case-referent study among Finnish woodworkers. *Scandinavian Journal Of Work, Environment & Health* 1985; 11: 409-415.
115. OECD. Formaldehyde. SIDS Initial Assessment Report. In Edition Paris: OECD, SIDS 2002.
116. Formaldehyde, 2-butoxyethanol and 1-tert-butoxypropan-2-ol. In IARC monographs on the evaluation of carcinogenic risks to humans. Vol. 88, Edition Lyon: IARC 2006.
117. Levine RJ, Andjelkovich DA, Shaw LK. The mortality of Ontario undertakers and a review of formaldehyde-related mortality studies. *Journal Of Occupational*

Medicine.: Official Publication Of The Industrial Medical Association 1984; 26: 740-746.

118. Nelson N, Levine RJ, Albert RE et al. Contribution of formaldehyde to respiratory cancer. *Environ Health Perspect* 1986; 70: 23-35.
119. Blair A, Saracci R, Stewart PA et al. Epidemiologic evidence on the relationship between formaldehyde exposure and cancer. *Scandinavian Journal Of Work, Environment & Health* 1990; 16: 381-393.
120. Partanen T. Formaldehyde exposure and respiratory cancer--a meta-analysis of the epidemiologic evidence. *Scandinavian Journal of Work, Environment & Health* 1993; 8-15.
121. Bosetti C, McLaughlin JK, Tarone RE et al. Formaldehyde and cancer risk: a quantitative review of cohort studies through 2006. *Annals of Oncology* 2008; 19: 29-43.
122. Banchand AM, Mundt KA, Mundt DJ, Montgomery RR. Epidemiological studies of formaldehyde exposure and risk of leukaemia and nasopharyngeal cancer: A meta-analysis. *Critical Reviews in Toxicology* 2010; 40: 85-100.
123. Binazzi A, Ferrante P, Marinaccio A. Occupational exposure and sinonasal cancer: a systematic review and meta-analysis. *BMC Cancer* 2015; 15: 1-17.
124. Partanen T. Formaldehyde exposure and respiratory cancer--a meta-analysis of the epidemiologic evidence. *Scand J Work Environ Health* 1993; 19: 8-15.
125. Bachand AM, Mundt KA, Mundt DJ, Montgomery RR. Epidemiological studies of formaldehyde exposure and risk of leukemia and nasopharyngeal cancer: A meta-analysis. *Critical Reviews in Toxicology* 2010; 40: 85-100.
126. Coggon D, Ntani G, Harris EC, Palmer KT. Upper Airway Cancer, Myeloid Leukemia, and Other Cancers in a Cohort of British Chemical Workers Exposed to Formaldehyde. *American Journal of Epidemiology* 2014.

127. Meyers AR, Pinkerton LE, Hein MJ. Cohort mortality study of garment industry workers exposed to formaldehyde: Update and internal comparisons. *American Journal of Industrial Medicine* 2013; 56: 1027-1039.
128. Stroup DF, Berlin JA, Morton SC et al. Meta-analysis of observational studies in epidemiology: a proposal for reporting. Meta-analysis Of Observational Studies in Epidemiology (MOOSE) group. *JAMA: The Journal Of The American Medical Association* 2000; 283: 2008-2012.
129. Wells GA, Shea B, O'Connell D et al. The Newcastle-Ottawa Scale (NOS) for assessing the quality of non-randomised studies in meta-analyses. In Edition 2014.
130. Collins JJ, Esmen NA, Hall TA. A review and meta-analysis of formaldehyde exposure and pancreatic cancer. *American Journal Of Industrial Medicine* 2001; 39: 336-345.
131. Collins JJ, Lineker GA. A review and meta-analysis of formaldehyde exposure and leukemia. *Regul Toxicol Pharmacol* 2004; 40: 81-91.
132. Hernberg S, Collan Y, Degerth R et al. Nasal cancer and occupational exposures. Preliminary report of a joint Nordic case-referent study. *Scandinavian Journal Of Work, Environment & Health* 1983; 9: 208-213.
133. Krieger N. Formaldehyde and nasal cancer mortality. *Can Med Assoc J* 1983; 128: 248-249.
134. Zhang R, Jiang X-z. A preliminary report on the analysis pf malignant mortality among anatomists and their assistants. *Occup Health Shanghai* 1988; 2: 89-90.
135. Band PR, Nhu DL, Fang R et al. Cohort mortality study of pulp and paper mill workers in British Colombia, Canada. *Am J Epidemiol* 1997; 146: 186-194.
136. Dubrow R, Wegman DH. Cancer and occupation in Massachusetts: a death certificate study. *Am J Ind Med* 1984; 6: 207-230.

137. Friedman GD, Ury HK. Screening for possible drug carcinogenicity: second report of findings. *Journal Of The National Cancer Institute* 1983; 71: 1165-1175.
138. Vaughan TL, Strader C, Davis S, Daling JR. Formaldehyde and cancers of the pharynx, sinus and nasal cavity: II. Residential exposures. *International Journal Of Cancer. Journal International Du Cancer* 1986; 38: 685-688.
139. Walrath J, Rogot E, Murray J, Blair A. Mortality patters among US veterans by occupation and smoking status. In Edition 1985.
140. Chiazze L, Jr, Watkins DK, Fryar C. Historical cohort mortality study of a continuous filament fiberglass manufacturing plant. I. White men. *Journal Of Occupational And Environmental Medicine / American College Of Occupational And Environmental Medicine* 1997; 39: 432-441.
141. Coggon D, Pannett B, Acheson ED. Use of job-exposure matrix in an occupational analysis of lung and bladder cancers on the basis of death certificates. *J Nat Can Institute* 1984; 72: 61-65.
142. Costantini AS, Paci E, Miliga L, Buiatti E. Cancer mortality among workers in the Tuscan tanning industry. *Br J Ind Med* 1989; 46: 384-388.
143. Fayerweather WE, Pell S, Bender JR. Case-control study of cancer deaths in Du Pont workers with potential exposure to formaldehyde. In Clary JJ, Gibson JE, Waritz RS (eds): *Formaldehyde: toxicology, epidemiology, mechanisms*, Edition New York: Marcel Dekker, Inc 1983; 47-125.
144. Liebling T, Rosenman KD, Pastides H et al. Cancer mortality among workers exposed to formaldehyde. *American Journal Of Industrial Medicine* 1984; 5: 423-428.

145. Marsh G, Stone R, Youk A, al. e. Mortality among United States rock wool and slag wook workers: 1989 update. *J Occup Health Saf Aust N Z* 1996; 12: 297-312.
146. Partanen T, Kauppinen T, Hernberg S et al. Formaldehyde exposure and respiratory cancer among woodworkers--an update. *Scandinavian Journal Of Work, Environment & Health* 1990; 16: 394-400.
147. Walrath J, Fraumeni JF, Jr. Cancer and other causes of death among embalmers. *Cancer Research* 1984; 44: 4638-4641.
148. Bond GG, Flores GH, Shellenberger RJ et al. Nested case-control study of lung cancer among chemical workers. *American Journal Of Epidemiology* 1986; 124: 53-66.
149. Walrath J, Fraumeni JF, Jr. Mortality patterns among embalmers. *International Journal Of Cancer. Journal International Du Cancer* 1983; 31: 407-411.
150. Higgins JP, Thompson SG, Deeks JJ, Altman DG. Measuring inconsistency in meta-analyses. *BMJ* 2003; 327.
151. Sterne JAC, Egger M. Funnel plots for detecting bias in meta-analysis. *Journal of Clinical Epidemiology* 54: 1046-1055.
152. Edling C, Jarvholm B, Andersson L, Axelson O. Mortality and cancer incidence among workers in an abrasive manufacturing industry. *British Journal Of Industrial Medicine* 1987; 44: 57-59.
153. Bertazzi PA, Pesatori A, Guercilena S et al. Carcinogenic risk for resin producers exposed to formaldehyde: extension of follow-up. *La Medicina Del Lavoro* 1989; 80: 111-122.
154. Andjelkovich DA, Janszen DB, Brown MH et al. Mortality of iron foundry workers: IV. Analysis of a subcohort exposed to formaldehyde. *Journal Of*

155. Harrington JM, Shannon HS. Mortality study of pathologists and medical laboratory technicians. *British Medical Journal* 1975; 4: 329-332.
156. Stroup NE, Blair A, Erikson GE. Brain cancer and other causes of death in anatomists. *Journal Of The National Cancer Institute* 1986; 77: 1217-1224.
157. Hall A, Harrington JM, Aw TC. Mortality study of British pathologists. *American Journal Of Industrial Medicine* 1991; 20: 83-89.
158. Roush GC, Schymura MJ, Stevenson JM, Holford TR. Time and age trends for sinonasal cancer in Connecticut incidence and US mortality rates. *Cancer* 1987; 60.
159. West S, Hildesheim A, Dosemeci M. Non-viral risk factors for nasopharyngeal carcinoma in the Philippines: results from a case-control study. *International Journal Of Cancer. Journal International Du Cancer* 1993; 55: 722-727.
160. Vaughan TL, Stewart PA, Teschke K et al. Occupational exposure to formaldehyde and wood dust and nasopharyngeal carcinoma. *Occupational And Environmental Medicine* 2000; 57: 376-384.
161. Olsen JH, Asnaes S. Formaldehyde and the risk of squamous cell carcinoma of the sinonasal cavities. *British Journal Of Industrial Medicine* 1986; 43: 769-774.
162. Luce D, Leclerc A, Bégin D et al. Sinonasal cancer and occupational exposures: a pooled analysis of 12 case-control studies. *Cancer Causes Control* 2002; 13.
163. Pesch B, Pierl CB, Gebel M et al. Occupational risks for adenocarcinoma of the nasal cavity and paranasal sinuses in the German wood industry. *Occupational and Environmental Medicine* 2008; 65: 191-196.

164. Gerin M, Siemiatycki J, Nadon L et al. Cancer risks due to occupational exposure to formaldehyde: results of multi-site case control study in Montreal. *Int J Cancer* 1989; 44: 53-58.
165. Brownson RC, Alavanja MCR, Chang JC. Occupational risk factors for lung cancer among nonsmoking women: a case-control study in Missouri (United States). *Cancer Causes & Control* 1993; 4: 449-454.
166. Stefani ED, Boffetta P, Brennan P et al. Occupational Exposures and Risk of Adenocarcinoma of the Lung in Uruguay. *Cancer Causes & Control* 2005; 16: 851-856.
167. Checkoway H, Ray RM, Lundin JI et al. Lung cancer and occupational exposures other than cotton dust and endotoxin among women textile workers in Shanghai, China. *Occupational and Environmental Medicine* 2011; 68: 425-429.
168. Mahboubi A, Koushik A, Siemiatycki J et al. Assessment of the effect of occupational exposure to formaldehyde on the risk of lung cancer in two Canadian population-based case-control studies. *Scandinavian Journal of Work, Environment & Health* 2013; 401-410.
169. Jensen OM, Andersen SK. Lung cancer risk from formaldehyde. *Lancet* 1982; 1: 913.
170. Milham Jr S. Occupational mortality in Washington State 1950-1979. In Edition Cincinnati, Ohio: US Department of Health and Human Sciences, Public Health and Human Services, Centers for Disease Control 1983.
171. Hayes RB, Blair A, Stewart PA et al. Mortality of U.S. embalmers and funeral directors. *American Journal Of Industrial Medicine* 1990; 18: 641-652.
172. Hansen J, Olsen JH. Formaldehyde and cancer morbidity among male employees in Denmark. *Cancer Causes & Control: CCC* 1995; 6: 354-360.

173. Stellman SD, Demers PA, Colin D, Boffetta P. Cancer mortality and wood dust exposure among participants in the American Cancer Society Cancer Prevention Study-II (CPS-II). *Am J Ind Med* 1998; 34: 229-237.
174. Siew SS, Kauppinen T, Kyyrönen P et al. Occupational exposure to wood dust and formaldehyde and risk of nasal, nasopharyngeal, and lung cancer among Finnish men. *Cancer Manag Res* 2012; 4.
175. Checkoway H, Dell LD, Boffetta P et al. Formaldehyde Exposure and Mortality Risks From Acute Myeloid Leukemia and Other Lymphohematopoietic Malignancies in the US National Cancer Institute Cohort Study of Workers in Formaldehyde Industries. *Journal of Occupational and Environmental Medicine* 2015; 57: 785-794.
176. Bosetti C, McLaughlin JK, Tarone RE et al. Formaldehyde and cancer risk: a quantitative review of cohort studies through 2006. *Ann Oncol* 2008; 19.
177. Kalinić N, Vadić V. The Adsorption of Formaldehyde on Suspended Particles investigated on model systems. *Environmental Monitoring and Assessment* 2000; 60: 15-23.
178. IARC. IARC classifies formaldehyde as carcinogenic to humans. Press release no. 133. In Edition 2004.
179. Kopec JA, Esdaile JM. Bias in case-control studies. A review. *Journal of Epidemiology and Community Health* 1990; 44: 179-186.
180. Blair A, Saracci R, Stewart PA et al. Epidemiologic evidence on the relationship between formaldehyde exposure and cancer. *Scand J Work Environ Health* 1990; 16.
181. Partanen T. Formaldehyde exposure and respiratory cancer: a meta-analysis of the epidemiologic evidence. *Scand J Work Environ Health* 1993; 19.

182. Bernard RM, Borokhovski E, Schmid RF, Tamim RM. An exploration of bias in meta-analysis: the case of technology integration research in higher education. *Journal of Computing in Higher Education* 2014; 26: 183-209.
183. Collins JJ, Acquavella JF. Review and meta-analysis of studies of acrylonitrile workers. *Scandinavian Journal Of Work, Environment & Health* 1998; 24: 71-80.
184. Hill AB. The Environment and Disease: Association or Causation? *Proceedings of the Royal Society of Medicine* 1965; 58: 295-300.
185. Collins JJ, Acquavella JF, Esmen NA. An updated meta-analysis of formaldehyde exposure and upper respiratory tract cancers. *J Occup Environ Med* 1997; 39.
186. Marsh GM, Stone RA, Henderson VL. A reanalysis of the National Cancer Institute study on lung cancer mortality among industrial workers exposed to formaldehyde. *J Occup Med* 1992; 34: 42-44.
187. Blair A, Stewart P, O'Berg M et al. Mortality among industrial workers exposed to formaldehyde. *Journal Of The National Cancer Institute* 1986; 76: 1071-1084.
188. Blair A, Stewart PA, Hoover RN. Mortality from lung cancer among workers employed in formaldehyde industries. *Am J Ind Med* 1990; 17: 683-699.
189. Blair A, Stewart PA. Comments on the Sterling and Weinkam analysis of data from the National Cancer Institute formaldehyde study. *Am J Ind Med* 1994; 25: 603-606.
190. Sterling TD, Weinkam JJ. Reanalysis of lung cancer mortality in a National Cancer Institute study on mortality among industrial workers exposed to formaldehyde. *J Occup Med* 1988; 30: 895-901.

191. Sterling TD, Weinkam JJ. Mortality from respiratory cancer (including lung cancer) among workers employed in formaldehyde industries. *Am J Ind Med* 1994; 25: 593-602.
192. Sterling TD, Weinkam JJ. Comments on the Blair and Stewart comments on the Sterling and Weinkam analysis of data from the National Cancer Institute formaldehyde study. *Am J Ind Med* 1995; 27: 301-305.
193. Blair A, Stewart PA. Comments on the reanalysis of the National Cancer Institute study of workers exposed to formaldehyde. *Journal Of Occupational Medicine.: Official Publication Of The Industrial Medical Association* 1989; 31: 881-884.
194. Sterling TD, Weinkam JJ. Reanalysis of lung cancer mortality in a National Cancer Institute Study on "Mortality among industrial workers exposed to formaldehyde". *Experimental Pathology* 1989; 37: 128-132.
195. Callas PW, Pastides H, Hosmer DW, Jr. Lung cancer mortality among workers in formaldehyde industries. *Journal Of Occupational And Environmental Medicine / American College Of Occupational And Environmental Medicine* 1996; 38: 747-748; discussion 749-751.
196. Collins JJ, Caporossi JC, Utidjian HMD. Formaldehyde exposure and nasopharyngeal cancer: re-examination of the National Cancer Institute study and an update of one plant. *J Nat Can Institute* 1988; 80: 376-377.
197. Collins JJ, Caporossi JC, Utidjian HHD. Formaldehyde exposure and nasopharyngeal cancer: reexamination of the National Cancer Institute study, and an update in one plant. *J Natl Can Inst* 1987; 78: 192.
198. Stayner L, Smith AB, Reeve G et al. Proportionate mortality study of workers in the garment industry exposed to formaldehyde. *American Journal Of Industrial Medicine* 1985; 7: 229-240.

199. Stayner LT, Elliott L, Blade L et al. A retrospective cohort mortality study of workers exposed to formaldehyde in the garment industry. *American Journal Of Industrial Medicine* 1988; 13: 667-681.
200. Pinkerton LE, Hein MJ, Stayner LT. Mortality among a cohort of garment workers exposed to formaldehyde: an update. *Occupational And Environmental Medicine* 2004; 61: 193-200.
201. Acheson ED, Barnes HR, Gardner MJ et al. Formaldehyde in the British chemical industry. An occupational cohort study. *Lancet* 1984; 1: 611-616.
202. Gardner MJ, Pannett B, Winter PD, Cruddas AM. A cohort study of workers exposed to formaldehyde in the British chemical industry: an update. *British Journal Of Industrial Medicine* 1993; 50: 827-834.
203. Acheson ED, Barnes HR, Gardner MJ et al. Formaldehyde process workers and lung cancer. *Lancet* 1984; 1: 1066-1067.
204. Acheson ED, Barnes HR, Gardner MJ et al. Cohort study of formaldehyde process workers. *Lancet* 1984; 2: 403.
205. Coggon D, Harris EC, Poole J, Palmer KT. Extended follow-up of a cohort of British chemical workers exposed to formaldehyde. *Journal Of The National Cancer Institute* 2003; 95: 1608-1615.
206. Bertazzi PA, Pesatori AC, Radice L et al. Exposure to formaldehyde and cancer mortality in a cohort of workers producing resins. *Scandinavian Journal Of Work, Environment & Health* 1986; 12: 461-468.
207. Andjelkovich DA, Mathew RM, Richardson RB, Levine RJ. Mortality of iron foundry workers: I. Overall findings. *Journal Of Occupational Medicine.*: Official Publication Of The Industrial Medical Association 1990; 32: 529-540.

208. Levine RJ, Andjelkovich DA, Shaw LK, Dalcors RD. Mortality of Ontario undertakers: a first report. In Formaldehyde: toxicology, epidemiology, mechanisms, Edition New York, NY: Marcel Dekker, Inc 1983; 127-146.
209. Stroup NE. A retrospective follow-up study of mortality in anatomists. In Edition New Haven, CT: Yale University 1984.
210. Harrington JM, Oakes D. Mortality study of British pathologists 1974-80. British Journal Of Industrial Medicine 1984; 41: 188-191.
211. Roush GC, Walrath J, Stayner LT et al. Nasopharyngeal cancer, sinonasal cancer, and occupations related to formaldehyde: a case-control study. Journal Of The National Cancer Institute 1987; 79: 1221-1224.
212. Hildesheim A, Dosemeci M, Chan CC et al. Occupational exposure to wood, formaldehyde, and solvents and risk of nasopharyngeal carcinoma. Cancer Epidemiology, Biomarkers & Prevention: a Publication Of The American Association For Cancer Research, Cosponsored By The American Society Of Preventive Oncology 2001; 10: 1145-1153.
213. Stellman SD, Boffetta P, Garfinkel L. Smoking habits of 800,000 American men and women in relation to their occupations. American Journal Of Industrial Medicine 1988; 13: 43-58.
214. Stellman SD, Garfinkel L. Cancer mortality among woodworkers. American Journal Of Industrial Medicine 1984; 5: 343-357.

APPENDIX 1 – SUMMARY OF THE EPIDEMIOLOGICAL STUDIES INCLUDED IN THE META-ANALYSIS IN CHAPTER 5

A1.1 THE NATIONAL CANCER INSTITUTE (NCI) COHORT

The NCI cohort study is a study of industrial workers employed at 10 formaldehyde-producing or formaldehyde-using plants in the USA. A series of publications have been produced of some of the individual plants and of all the plants combined. They are summarised below.

The latest analysis of the full NCI cohort, which updated earlier publications of analyses of the same cohort or parts thereof [89, 186-197], included 25,619 workers employed at one of ten plants prior to 1 January 1966 in the US formaldehyde industry [82].

Subjects were followed-up from the year in which employment records were considered complete at the plant - earliest 1934 - or date of first employment at the plant, if later, to the end of 204. Exposures to formaldehyde were estimated from work histories up to 1980 and were based on job title, tasks, and visits to the plant by industrial hygienists, discussions with plant workers and managers, and monitoring data. Exposure assessments over time to 1980 were made for cumulative exposure in ppm-years, average exposure intensity in ppm, duration of exposure in years, highest peak exposure category in ppm, exposure to formaldehyde-containing particulates summarised as ever/never, duration of exposure to 11 other suspected carcinogens and widely-used chemicals, and duration of working as a chemist or laboratory technician in years. All exposures were lagged by 15 years to account for the latency of solid cancers. SMRs were calculated using the person-years method and relative risks were calculated using Poisson regression models, adjusting for calendar year, age, sex, race and pay category.

Potential confounding by other chemicals or work as a chemist or laboratory technician was evaluated.

Based on comparisons with the US population, mortality from nasopharyngeal cancer was non-significantly raised among those exposed to formaldehyde (SMR = 184, 95% CI 84 to 349, 9 deaths). There was a highly statistically significant increasing trend in relative risk for increased level of peak exposure ($p < 0.001$). The increasing trends for nasopharyngeal cancer for average exposure intensity ($p = 0.09$) and cumulative exposure ($p = 0.06$) were of borderline. There was an increased RR of nasopharyngeal cancer in the highest categories of formaldehyde exposure, with eight of 10 deaths occurring among formaldehyde-exposed workers. The RR was 7.66 (0.94 to 62.34) for peak exposures, 11.54 (1.38 to 96.81) for average exposure intensity and 2.94 (0.65 to 13.28) for cumulative exposure, compared with the lowest exposure category. Duration of exposure, regardless of exposure level, resulted in a RR of 2.53 (0.4 to 15.0) in the highest category (>15) years, p -trend = 0.4. Five of the 10 deaths occurred in the plant singled out for attention by Marsh in his critique of the NCI study. They conducted an influence analysis excluding one plant at a time. When any plant other than plant 1 was excluded, the results were similar to the overall analysis. When plant 1 was excluded, the number of nasopharyngeal cancer deaths was two in the highest peak exposure category RR = 3.36 (0.3 to 37.27), one in the highest average intensity category and zero in the highest cumulative exposure category.

Mortality from cancer of the nose and nasal cavity was around expected SMR = 90 (18 to 262, 3 deaths). There was no evidence of any exposure-response relationship for nose and nasal sinus cancer.

There were 1,291 deaths from lung cancer and a significantly elevated SMR in those exposed to formaldehyde 114 (1.07 to 1.20). [82]In internal analyses, there was a

significantly decreased risk in the highest category of peak RR = 0.77 (0.67 to 0.90, p-trend = 0.001) and cumulative formaldehyde exposure RR = 0.78 (0.66 to 0.93, p-trend = 0.02).

A1.2 THE NIOSH COHORT

The latest analysis of the NIOSH cohort [127] updated two earlier analyses of the same cohort [198-200] and included 11,043 workers, predominantly white and female, exposed to formaldehyde for three months or more in three garment manufacturing plants. Follow-up was from 1955 to 2008. Forty two per cent of the cohort was first exposed before 1963, when formaldehyde exposure levels were thought to be higher than in later years. The median duration of exposure was 3.3 years and the median time since first exposure was 39.4 years. The authors reported that there was no evidence of the presence of other potentially carcinogenic exposures at the three plants.

Mortality in the cohort was compared with the US and state populations using SMRs. Stratified analyses by duration of exposure, time since first exposure and year of first exposure were undertaken, together with tests for trends in SMRs. Multiple cause analyses utilising non-underlying causes of death were also employed. Poisson regression analysis was used to examine rate ratios by exposure duration.

No deaths from cancer of the nasopharynx, SMR = 0 (0 to 277), or the nose and nasal sinuses, SMR = 0 (0 to 389), were observed. Mortality from lung cancer was similar to that expected based on US rates with SMR = 104 (95% CI 92 to 117, 267 deaths). The results using state rates were similar. Lung cancer was elevated among workers first exposed in 1971 or later and was highest among person-time with < 10 years since first exposure. There was no clear pattern of increased risk across increasing exposure duration categories and was less than expected for the longest duration of exposure (10+ years) category compared to the US population. Overall, although lung cancer risk was

elevated in some subgroups, there was little evidence that formaldehyde was associated with lung cancer.

A1.3 THE BRITISH COHORT

The latest analysis of the British cohort [126] updated earlier analyses of the same cohort [201-205]. The cohort comprised 14,008 men who had been employed at six British chemical factories from 1938 to 1964 at a time when formaldehyde was produced or used, and follow-up was from 1941 to 2012. Occupational histories were abstracted from employment records. No formaldehyde measurements were available before 1970, but later measurements and workers' recall of irritant symptoms resulted in the following exposure classification: background (< 1 ppm); low (0.1 - 0.5 ppm); moderate (0.6 - 2.0 ppm); high (> 2.0 ppm); and unknown. Some of the exposures may have been through inhalation of formaldehyde particles or particles of formaldehyde-based products. In addition to formaldehyde, other potential carcinogenic exposures were present although any exposures were deemed to have been relatively low. SMRs were calculated using national mortality rates and national rates with an adjustment for locality. Poisson regression analysis was used to test for trends across ordered categorical exposures.

There was only one death from nasopharyngeal cancer, with 2.0 expected (SMR = 50, 95% CI 13 to 279) based on national rates, and the man concerned had not worked in a job with high exposure to formaldehyde.

Overall there was an elevated risk of lung cancer (813 observed v 645.8 expected), which was significantly elevated. When broken down by factory, one factory had no elevated risk, but the risk was elevated at each of the other five factories. For nose and nasal sinuses there was 0 deaths v 0.9 expected and for nasopharyngeal cancer, the only

death occurred in a man with moderate/low exposure (1.7 deaths expected for exposures above background).

For lung cancer, there was no exposure-response relationship. Indeed risk was lower in men with prolonged high exposure than in those highly exposed for less than one year and this pattern persisted when each man's first 35 years of follow-up were disregarded. The authors suggested that their finding for lung cancer was likely explained by non-occupational factors.

A1.4 OTHER INDUSTRIAL COHORT STUDIES

The cancer morbidity and mortality of a cohort of 521 men manufacturing abrasive materials has been examined [152]. The cohort was followed-up from 1958 to 1983 and included men who were employed at least five years some time between 1955 and 1983. During the manufacture of grinding wheels bound by formaldehyde resins, there was moderate exposure to formaldehyde of 0.1 to 1 mg/m³. During the manufacture of abrasive belts, workers had an intermittent heavy exposure to formaldehyde of 20 to 30 mg/m³. Mortality and morbidity were compared with the national Swedish data. There was one case of nasopharyngeal cancer and no cases of sinonasal cancer.

The most recent update of an Italian cohort of workers exposed to formaldehyde during the manufacture of resins [153] extended the follow-up of an earlier publication [206]. From 1974 to 1979 the mean concentration ranged from 0.064 ppm to 2.13 ppm and the peak exposure ranged from 0.33 ppm to 6.53 ppm. The latest analysis included 1330 male workers who worked at the factory for 30 days or more between 1959 and 1980 and who were followed-up until 1986. Local and national populations were used as the basis for calculating SMRs. In the latest report, no mention is made of nasopharyngeal cancer [153]. In the earlier report, there were no cases of nasal cancer, with 0.0327

expected [206]. The SMR for lung cancer for those exposed to formaldehyde based on local rates was 69 (95% CI 25 to 150, 6 deaths) [153].

The mortality experience of a subcohort 3929 men who worked at an iron foundry in the US and who were exposed to formaldehyde at for at least six months from 1960 to 1987 was compared with the national population via SMRs and lung cancer mortality was also compared with an internal group of 2032 men unexposed to formaldehyde via Poisson regression analysis [154]. Findings for the full cohort have previously been reported [207]. Follow-up in this report was from 1950 to 1989 [154]. According to NIOSH surveys, the level of formaldehyde found in iron foundries ranged from < 0.02 to 18.3 ppm. Each of 107 occupational titles was categorised as high (1.5 ppm), medium (0.55 ppm), low (0.05 ppm) or no exposure to formaldehyde. Among workers exposed to formaldehyde the SMR for cancer of the buccal cavity and pharynx was 101 (95% CI 48 to 286, 6 deaths) and among the unexposed, the SMR was 169 (95% CI 54 to 395, 5 deaths). There was only one death from nasopharyngeal cancer and this was in an unexposed worker. An approximate 95% confidence interval for the nasopharyngeal cancer SMR in formaldehyde-exposed men based on the expected number of lung cancer deaths is 0 to 1816. There were no deaths from nasal cancer; an equivalent approximate 95% CI for the SMR is 0 to 7306. For exposed men, the SMR for lung cancer was 120 (95% CI 89 to 158, 51 deaths) and for men unexposed to formaldehyde was 119 (95% CI 84 to 163, 38 deaths). From the Poisson regression, the rate ratio for lung cancer for exposed versus unexposed was 0.71 (95% CI 0.44 to 1.21) [154].

A1.5 COHORT STUDIES OF PATHOLOGISTS, ANATOMISTS, FUNERAL DIRECTORS AND EMBALMERS

The mortality of 2079 British pathologists alive in 1955 was followed-up to 1983.

SMRs were calculated on the basis of rates for Great Britain. Five unspecified malignant neoplasms were included in the total, and it is unclear whether any of them were nasopharyngeal or nasal cancers. The SMR for lung cancer was 39 (95% CI 20 to 71, 11 deaths) [155].

A cohort study of 1477 male undertakers in Ontario, Canada, examined mortality of those first licensed from 1928 to 1957 and followed up from 1950 to 1977 [117]. This study updated a slightly earlier report that used US rates for the comparison [208].

SMRs for the latest report were calculated using mortality rates for Ontario. The SMR for buccal cavity and pharynx cancer was 48 (95% CI 1 to 265, 1 death); it is not clear whether or not this death was due to nasopharyngeal cancer. There were no deaths due to nasal cancer (95% CI 0 to 1498) and the SMR for lung cancer was 94 (95% CI 57 to 147, 19 deaths) [117].

A retrospective cohort has been established of 2317 male anatomists in the US who joined the American Association of Anatomists between 1888 and 1969 and who were living in the US when they joined [156], with the work being completed for a PhD thesis [209]. Follow-up was from 1925 to 1979 and SMRs were calculated on the basis of US rates. The SMR for buccal cavity and pharynx cancer was 20 (95% CI 0 to 80, 1 death) and the death was not due to nasopharyngeal cancer. The SMR for cancer of the nasal cavity and sinuses was 0 (95% CI 0 to 720, 0 deaths). The SMR for lung cancer was 30 (95% CI 10 to 50, 12 deaths).

A mortality experience of a separate cohort of 3,872 British pathologists has also been examined [157], which updated an earlier report [210]. Members of the Royal College

of Pathologists from 1973 to 1986 who were alive at the end of 1973 were followed up from 1974 to 1987. SMRs were calculated on the basis of rates for England and Wales and for Scotland. Four unspecified neoplasms were included in the total for England and Wales in the latest update, and it is unclear whether any of them were nasopharyngeal or nasal cancers. The SMR for lung cancer for England and Wales was 19 (95% CI 9 to 36, 9 deaths), and for Scotland the SMR was 60 (95% CI 16 to 154, 4 deaths) [157].

A1.6 CASE-CONTROL STUDIES OF NASOPHARYNGEAL CANCER

A population-based case-control study of male nasopharyngeal cancer cases ($n = 173$) was carried out with subjects identified from the Connecticut Tumour Registry from 1935 to 1975 [211]. Controls ($n = 605$) were deaths among Connecticut residents during the same time period and were selected randomly without stratification or matching. Exposure to formaldehyde was assessed according to a probability and level of exposure blind to case-control status in the same way as the NCI cohort [187], with occupational histories being constructed from death certificates and City Directories. Odds ratios adjusted for age at death, year at death, and availability of occupational information were calculated using logistic regression, but were not adjusted for other occupational exposures. For nasopharyngeal cancer the OR for probably exposed to some level for most of working life compared with all others was 1.0 (95% CI 0.7 to 2.4). For probable exposure to some level for most of working life and probable exposure to some level 20 or more years prior to death, the OR was 1.3 (95% CI 0.7 to 2.4). For probably exposed to some level for most of working life and probably exposed to high level in some year the OR was 1.4 (95% CI 0.6 to 3.1). For those probably exposed to some level for most of working life and probably exceeded high level 20 years or more before death, the OR was 2.3 (95% CI 0.9 to 6.0).

A case-control study of nasopharyngeal cancer was carried out in the Philippines on 104 predominantly non-Chinese cases and 205 hospital and community controls [159]. Risk factor information was obtained through personal interview. The occupational history of each subject was reviewed in the absence of knowledge of case-control status by an industrial hygienist. Exposure to formaldehyde, solvents, dusts, exhaust and pesticides were assessed. After controlling for confounding, subjects who were first exposed for formaldehyde 25 years or more prior to diagnosis or interview or who first exposed before the age of 25 were found, in relation to those never exposed, to be at a 4.0 fold excess risk of disease (1.4 to 12.3). The equivalent risk for <25 years since first exposure was 1.2 (0.41 to 3.6). The adjusted odds ratio for <15 years duration of exposure was 2.7 (1.1 to 6.6) and for 15 years or more was 1.2 (0.48 to 3.2).

During 1990-1992, 282 Chinese residents in Malaysia with histologically confirmed nasopharyngeal carcinoma were interviewed about occupational history, diet, alcohol consumption, and tobacco use, as were an equal number Malaysian Chinese population controls, pair-matched to cases by age and sex [98]. Exposure to 20 kinds of workplace substances, solar and industrial heat, and cigarette smoking were analysed univariably and multivariably. No significant crude or adjusted association was found between nasopharyngeal carcinoma and formaldehyde, adjusted OR – 0.71 (0.34 to 1.43).

A multi-centred population-based case-control study was carried out at five cancer registries in the United States [160]. Cases (n = 196) with a newly diagnosed nasopharyngeal cancer between 1987 and 1993 and controls (n = 244) selected over the same period from the general population through random-digit dialing participated in a structured telephone interview which inquired about suspected risk factors for the disease, including lifetime history of occupational and chemical exposure. Potential exposure to formaldehyde was assessed blind to case-control status on a job by job basis by experienced industrial hygienists. For formaldehyde, after adjusting for cigarette

use, race, and other risk factors, a trend of increasing risk of squamous and unspecified epithelial carcinoma was found for increase duration ($p = 0.014$) and cumulative exposure ($p = 0.033$) but not for maximum exposure concentration. The odds ratio for people cumulatively exposed to >1.10 ppm-years was 3.0 (1.3 to 6.6) compared with those considered unexposed. In analyses limited to jobs considered definitely exposed, these trends became stronger. The associations were most evidence among cigarette smokers. By contrast, there was no association between potential exposure to formaldehyde and undifferentiated and non-keratinising carcinomas.

A case-control study was conducted among 375 newly diagnosed cases of nasopharyngeal carcinoma in Taipei, Taiwan and 325 community controls matched to cases on age, sex and geographical residence. Most cases were diagnosed with non-keratinising and undifferentiated carcinomas [212]. A complete occupational history was obtained via a personal interview and blindly assessed by an industrial hygienist for intensity and probability of exposure to wood dust, formaldehyde and solvents. Information on socio-demographic characteristics, cigarette smoking, dietary consumption of nitrosamines, and other potential confounding factors was obtained via a personal interview. Blood specimens were tested for human antigen class I/II genotypes, polymorphisms in cytochrome P450 2E1 genotype, and various anti-EBV antibodies known to be associated with nasopharyngeal carcinoma. Individual exposed to formaldehyde had a non-significant increased risk $RR = 1.4$ (0.93 to 2.2). Those exposed to formaldehyde for >10 years had an adjusted RR of 1.6 (0.91 to 2.9). The association between formaldehyde and nasopharyngeal carcinoma was stronger in analyses restricted EBV seropositive individuals $RR = 2.7$ (1.2 to 5.9). However, no dose response was observed with increasing duration or cumulative use.

A1.7 CASE-CONTROL STUDIES OF SINONASAL CANCER

A study of 759 histologically verified cancer of the nasal cavity (287 cases) and paranasal sinuses (179 cases) and 2,465 cancer controls diagnosed in Denmark between 1970 and 1982 was conducted to investigate the importance of occupational exposure to formaldehyde [161]. Information on job history for cases and controls was derived from a national data linkage system and exposure to formaldehyde and wood dust was assessed by industrial hygienists unaware of the case-control status of the patients. The exposure rates for formaldehyde among male and female controls were 4.2% and 0.1% respectively. After proper adjustment for contemporary wood dust exposure, relative risks of 2.3 (0.9 to 5.8) for squamous cell carcinoma and 2.2 (0.7 to 7.2) for adenocarcinoma of the nasal cavity and paranasal sinuses were detected among men who had ever been exposed to formaldehyde in their job compared with those never exposed. The introduction of 10 year latency did not change the risk estimates substantially.

A population-based case-control study of male sinonasal cancer (n = 198) cases was carried out with subjects identified from the Connecticut Tumour Registry from 1935 to 1975 [211]. Controls (n = 605) were deaths among Connecticut residents during the same time period and were selected randomly without stratification or matching. Exposure to formaldehyde was assessed according to a probability and level of exposure blind to case-control status in the same way as the NCI cohort [187], with occupational histories being constructed from death certificates and City Directories. Odds ratios adjusted for age at death, year at death, and availability of occupational information were calculated using logistic regression, but were not adjusted for other occupational exposures. For sinonasal cancer the OR for probably exposed to some level for most of working life compared with all others was 0.8 (95% CI 0.5 to 1.3). For probable exposure to some level for most of working life and probable exposure to

some level 20 or more years prior to death, the OR was 1.0 (95% CI 0.5 to 1.8). For probably exposed to some level for most of working life and probably exposed to high level in some year the OR was 1.0 (95% CI 0.5 to 2.2). For those probably exposed to some level for most of working life and probably exceeded high level 20 years or more before death, the OR was 1.5 (95% CI 0.6 to 3.9).

Data from 12 case-control studies of sinonasal cancer were pooled and re-analysed [162]. The pooled data set included 195 adenocarcinoma cases (169 men, 26 women), 432 squamous cell carcinomas (330 men, 102 women) and 3,136 controls (2,349 men, 787 women). Occupational exposure to formaldehyde, silica dust, textile dust, coal dust, flour dust, asbestos and man-made mineral fibres were assessed with a job-exposure matrix. Odds ratios were adjusted for age, study, wood dust, and leather dust, or other occupational exposures when relevant. A significantly increase risk of adenocarcinoma was associated with exposure to formaldehyde. The ORs for the highest level of exposure were 3.0 (1.1 to 5.7) among men and 6.2 (2.0 to 19.7) among women. An elevated risk of squamous cell carcinoma was observed among men OR = 2.5 (0.6 to 10.1) and women OR = 3.5 (1.2 to 10.5) with a high probability of exposure to formaldehyde.

An industry-based case-control study with 86 male adenocarcinomas of the nasal cavity and paranasal sinuses and 204 controls was conducted in the German wood working industries[163]. Cumulative and average wood-dust exposure was quantified with a job-exposure matrix. Probability of exposure to formaldehyde and other relevant exposures were semi-quantitatively rated. Logistic regression analyses were carried out conditional on age and adjusted for smoking and other factors. No significant associations were estimated for formaldehyde.

A1.8 CASE-CONTROL STUDIES OF LUNG CANCER

A nested case-control study of lung cancer among workers at a US chemical plant was carried out [148]. The cohort consisted of 19,608 males employees. Included in the case-control study were 308 lung cancer deaths observed between 1940 and 1981. Two control groups, one a decedent and other a “living” series, were individually matched to cases one-for-one. Interviews were conducted with subjects or their next of kin to collect information on smoking and other potential confounders. These data were combined with employee work history records and industrial hygiene data to form the basis of the analyses. Traditional stratification methods and conditional logistic regression were employed to examine for effect modification and to control for confounding. The relative risk for formaldehyde was 0.62 (0.28 to 1.34), based on nine exposed cases. When restricted to exposures 15 or more years prior to death of the cases, the RR was 0.31 (0.11 to 0.86), based on four cases.

A case-control study was undertaken in Montreal to investigation the possible associations between occupational exposures and a number of cancers including lung cancer.[164]. In total, 3,726 cancer patients and 533 population controls were interviewed to obtain detailed lifetime job histories and information on potential confounders. Each job history was translated into a history of occupational exposures. For short exposure, the OR was 1.0 (0.6 to 1.8, 62 exposed cases). For long-low exposure, the OR was 0.5 (0.3 to 1.0, 39). For long-medium exposure, the OR was 1.0 (0.6 to 1.7, 50). For long-high exposure, the OR was 1.0 (0.4 to 2.4, 24). Thus there was no persuasive evidence of an increased risk of any type of cancer among men exposed to various levels of formaldehyde, but the possibility of a small increase in risk could not be ruled out.

Respiratory cancer was examined in relation to occupational formaldehyde exposure in a case-referent study (136 cases, 408 referents) nested in a woodworker cohort [146]. Plant- and time-specific job-exposure matrices were constructed for formaldehyde exposure. Over 3 ppm-months of formaldehyde exposure was associated with an OR of 1.4 (90% CI: 0.5 to 4.1). The odds ratio for lung cancer was near unity, the excess risk concentrating on the upper respiratory tract. That for combined exposure to formaldehyde-phenol exposure (all respiratory cancers) was 1.6 (90% CIL 0.6 to 4.4) but 1.0 for formaldehyde only.

A population-based case-control study was carried out in Missouri in the United States [165]. Incident cases of lung cancer ($n = 429$) were identified through the Missouri cancer registry for the period 1986 to 1991 and included 294 lifelong non-smokers and 135 former smokers who had stopped at least 15 years prior to diagnosis or had smoked for less than one pack-year. Controls ($n = 1,021$) were selected through drivers' license and Medicare files. The odds ratio for exposure to formaldehyde, adjusted for age and history of previous lung disease was 0.9 (0.2 to 3.3) for all subjects and 0.9 (0.2 to 3.3) among lifetime non-smokers.

A case-control study of men at four major hospitals in Montevideo was conducted to investigate risks of lung adenocarcinoma [166]. Interviews were carried out among 339 cases and 1,014 hospital-based controls. Odds ratios were calculated after adjustment for tobacco smoking, residence, urban/rural status, education, body mass index. The odds ratio of ever exposed to formaldehyde was 1.7 (1.1 to 2.8). The odds ratio for exposure 1-20 years compared with not smoking was 0.9 (0.4 to 1.9) and for 21 years or more was 3.0 (1.6 to 5.8) yielding a statistically significant positive trend ($p = 0.004$).

A study to examine the effect of formaldehyde and other occupational exposures on lung cancer risk among 267,400 female textile workers in Shanghai, China was carried

out [167]. A case-cohort study nested within the cohort was used to compare work assignments and exposure histories of 628 incident lung cancer cases, diagnosed during 1989-1998, with those of a reference subcohort of 3,188 workers. Exposures were reconstructed with a job-exposure matrix developed specifically for textile factories. Cox proportional hazards models were applied to estimate age/smoking-adjusted relative risks (hazard ratios) and risk gradients associated with job assignments and specific agents other than cotton dust and endotoxin (such as formaldehyde). Increased risk, although statistically imprecise, were noted for 10 years or more exposure to formaldehyde, adjusted hazards ratio 2.1 (0.4 to 11.0).

Data were collected in two population-based case-control studies conducted in Montreal, Canada. Cases were individuals diagnosed with incident, histologically confirmed lung cancer [168]. Controls were randomly selected from electoral lists and frequency-matched to cases by age, sex and electoral district of residence. Interviews for the two studies were conducted in 1979-1986 and 1996-2002, using a virtually identical questionnaire to obtain lifetime occupational and smoking history and several lifestyle covariates. Experts reviewed the detailed work history for each participant to assess exposure to several occupational agents, including formaldehyde. Logistic regression was used to estimate odds ratios for the association between several metrics of formaldehyde exposure and lung cancer, adjusting for smoking and occupational and sociodemographic factors. In all, 2060 lung cancer and 2,046 population controls were interviewed and assessed for exposure. About 25% of subjects had been occupationally exposed to formaldehyde. The adjusted odds ratio for lung cancer was 1.06 (0.89 to 1.27) comparing ever versus never exposed to formaldehyde. Analyses for age at first exposure, average, and peak intensity of exposure also suggested an absence of association between formaldehyde exposure and lung cancer risk. Results did not vary by sex, lifetime smoking intensity or histological subtype.

A1.9 OTHER STUDIES INCLUDING REGISTRY-BASED

Data from the Danish cancer registry for the period 1943-1976 were used to examine the association with formaldehyde exposure in a total of 84 cases of lung cancer (79 male, 5 female) among Danish physicians [169]. Information on specialization and places of work during the professional career of these physicians with lung cancer was compared with the information for 252 controls, who were also physicians chosen to match cases for age, sex and survival at least until the time of lung cancer development. No male lung cancer cases had specialized in pathology (including forensic medicine and anatomy), and the risk in other medical specialties did not differ significantly from the risk among general practitioners. The lung cancer risk associated with employment at some time during the professional career in pathology, forensic medicine, or anatomy was not increased, odds ratio 1.0 (0.4 to 2.4).

A proportionate mortality study of occupational mortality among white residents of Washington State in the US examined 429,926 males deaths from 1950-1979 and 25,066 female deaths from 1974-1979 [170]. The PMR among funeral directors was 0.75 (0.40 to 1.28).

The mortality pattern of 1132 white male embalmers licensed to practice in New York State between 1902 and 1980 and who were known to have died between 1925 and 1980 has been examined [149]. PMRs and Proportionate Cancer Mortality Ratios (PCMRs) were calculated. There were no deaths from nasopharyngeal cancer with approximately 0.26 expected. There were also no deaths from nasal cancer with approximately 0.07 expected. The PMR for lung and pleural cancer was 108 (84 to 136, 72 deaths) and the PCMR for lung cancer was 111 (87 to 140, 70 deaths). The authors assumed that persons licensed only as embalmers experienced a greater cumulative exposure to formaldehyde than did embalmers who were also funeral

directors. The PMR for all respiratory system cancers for embalmers only was 94 (62 to 137, 27 deaths) and for those both embalmers and funeral directors the PMR was 112 (82 to 148, 47 deaths).

The mortality pattern of 1007 white male embalmers who were first licensed to practice in California from 1916 to 1978 and who were known to have died between 1925 and 1980 was examined [147]. PMRs and PCMRs were calculated. There were no deaths from nasopharyngeal cancer included, with approximately 0.17 expected. There was no nasal cancer cases with 0.6 expected. The PMR for lung and pleura cancer was 96 (69 to 130, 41 deaths) and the PCMR was 87 (62 to 118).

The mortality of 3649 white and 397 non-white male US embalmers and funeral directors in 32 states and the District of Columbia, who had died between 1975 and 1985 was examined in a proportional mortality study [171]. Deaths included in the Californian [149] and New York [147] studies were excluded. PMRs were calculated using US rates. The PMRs for nasopharynx were 189 (39 to 548, 3 deaths) among whites and 400 (10 to 2229, 1 death) among non-whites. For sinonasal cancer the PMR was 0 (0 to 200, 0 deaths) among whites and 0 (0 to 1872, 0 deaths) among non-whites. For lung cancer the SMR was 97 (86 to 109, 285 deaths) among whites and 75 (47 to 113, 23 deaths) among non-whites.

The Danish Cancer Registry was used to identify 126,347 men with cancer born between 1897 and 1964 whose cancer was diagnosed in the period 1970 to 1984 [172]. Individual employment histories were constructed via linkage with the Supplementary Pension Fund. All companies associated with the use or manufacture of more than one kg of formaldehyde were identified from the Danish Product Register, and thus a total of 2,041 employees were identified as having their longest work experience working in one of 265 companies, at least 10 years before diagnosis, involved in the use or

manufacture of formaldehyde. Workers were classified as probably exposed to low levels of formaldehyde (blue collar workers), as exposed to formaldehyde in the absence of wood-dust (blue collar workers), co-exposed to formaldehyde and wood-dust, and unknown in relation to formaldehyde exposure. Standardised Proportional Incidence Ratios (SPIRs) were calculated for those exposed to formaldehyde at least 10 years prior to diagnosis. (The SPIR is a measure of the proportion of cases of a defined cancer in the formaldehyde-associated companies relative to the proportion of cases of the same type of cancer among all employees in Denmark, adjusted for age and calendar time [172]). The SPIR for nasopharyngeal cancer was 1.3 (0.3 to 3.2, 4 deaths), for nasal cancer was 2.3 (1.3 to 4.0, 13 deaths) and for lung cancer was 1.0 (0.9 to 1.1, 410 deaths). For the group of workers classified as exposed to formaldehyde in the absence of wood-dust the SPIR for nasal cancer was 3.0 (1.4 to 5.7, 9 deaths) and for lung cancer was 1.0 (0.9 to 1.1, 250 deaths).

The most recent report of the cohort of 363,823 men enrolled in the American Cancer Society's Cancer Prevention Study [173] updated earlier reports on the cohort [213, 214]. The subcohort was a cross-sectional cohort that was assembled in 1982 and followed-up for six years until 1988. Three hundred and eighty seven subjects reported exposure to formaldehyde and 305 reported exposure to both wood dust and formaldehyde. Relative risks were estimated by incidence density ratios, relative to participants who did not report either employment in a wood occupation or regular exposure to wood dust, and were calculated using maximum likelihood with an adjustment for age and smoking status. The relative risk of lung cancer for men exposed to formaldehyde in the absence of a wood-related occupation was 0.93 (0.73 to 1.18, 104 deaths), and for those who had been in a wood-related occupation, the relative risk was 2.63 (1.25 to 5.51, 7 deaths).

The cohort of all Finnish men born between the years 1906 and 1945 and in employment during 1970 was followed up through the Finnish cancer registry for cases of cancer of the nose ($n = 292$), nasopharynx ($n = 149$), and lung ($n = 30.137$) during the period 1971 to 1995 [174]. The subjects' occupations, as recorded in the population census in 1970, were converted to estimates of exposure to wood dust, formaldehyde, asbestos and silica dust through the Finnish job-exposure matrix. Cumulative exposure was calculated based on the prevalence, average level, and estimated duration of exposure. The relative risk estimate for the cumulative exposure category were obtained by Poisson regression, with adjustment for smoking, socioeconomic status, and exposure to asbestos and/or silica dust. Workers exposed to formaldehyde had a RR of 1.18 (1.12 to 1.25) for lung cancer. There was no indication that cumulative exposure to formaldehyde would increase the risk of nasopharyngeal cancer. The authors thought that the slight excess risk for lung cancer might be due to residual confounding from smoking.